

# MANUAL DE MONITOREO Y EVALUACIÓN

*FOCUS on Young Adults*



FOCUS on Young Adults

**SERIE OPS/FNUAP NO. 3**

In July 2011, FHI became FHI 360.



FHI 360 is a nonprofit human development organization dedicated to improving lives in lasting ways by advancing integrated, locally driven solutions. Our staff includes experts in health, education, nutrition, environment, economic development, civil society, gender, youth, research and technology – creating a unique mix of capabilities to address today's interrelated development challenges. FHI 360 serves more than 60 countries, all 50 U.S. states and all U.S. territories.

Visit us at [www.fhi360.org](http://www.fhi360.org).

Los autores quieren agradecer la contribución de tantas personas que ayudaron en el desarrollo y la revisión de esta guía:

Queremos reconocer los esfuerzos de varios graduados asistentes en investigación que trabajaron con *FOCUS on Young Adults* en la Escuela de Salud Pública y Medicina Tropical del Departamento de Salud y Desarrollo Internacional en la Universidad de Tulane. Stephanie Mullen comenzó el proyecto compilando indicadores de programa. Gwendolyn Morgan preparó los anexos detallando referencias recomendadas y páginas web, y preparó las fórmulas para las tablas de indicadores. Emily Zielinski ayudó a preparar las tablas de indicadores y los anexos.

Nuestras colegas de FOCUS, Sharon Epstein, Lindsay Stewart, Barbara Seligman y Lisa Weiss leyeron las primeras versiones de esta guía y ofrecieron sugerencias. Sus comentarios nos ayudaron a siempre tener en cuenta a los trabajadores de programa que esperamos encuentren útil este documento.

Estamos agradecidos a Christine Steven, que forma parte del equipo de FOCUS, por su revisión técnica y las recomendaciones que hizo para reorganizar varios capítulos de la guía. También agradecemos la colaboración de Laura Sedlock, cuya edición clarificó los conceptos y unió los estilos de cada autor.

Las ideas y los conceptos que formaron la base de esta guía se discutieron en un Taller de Investigación y Evaluación de FOCUS en abril de 1998. Entre los participantes estaban Lisanne Brown (Universidad de Tulane), Nicola Bull (UNICEF), James Chui (FNUAP), Richard Colombia (*Pathfinder International*), Bruce Dick (UNICEF), Jane Ferguson (Organización Mundial de la Salud), Alix Grubel (*John Snow International*), Paula Hollerbach (*Academy for Educational Development*), Marge Horn (USAID), Merita Irby (*International Youth Foundation*), Lily Kak (CEDPA), Rebecka Lundgren (*Georgetown Institute for Reproductive Health*), Matilde Maddaleno (Organización Panamericana de la Salud), Leo Morris (*Centers for Disease Control*), Lisa Mueller (*John Snow International*), Ian Tweedie (*Johns Hopkins University Center for Communications Programs*), Stephanie Mullen (Universidad de Tulane), Phyllis Scattergood (*Education Development Center, Inc.*), Annetta Seecharan (*International Youth Foundation*), Linda Sussman (USAID), Katherine Weaver (Organización Panamericana de la Salud), Ellen Weiss (*Population Council/Horizons*) y Anne Wilson (PATH).

Contribuyeron comentarios y retroalimentación durante la etapa de revisión técnica Jane Bertrand (Universidad de Tulane), Carlos Brambila (*Population Council*, México), Eunyong Chung (USAID), Charlotte Colvin (*The Futures Group International*), Shanti Conly (USAID), Barbara de Zalduondo (USAID), Joyce Djaelani (PATH Indonesia), Maricela Dura (Fundación Mexicana para la Planeación Familiar), Natalia Espinoza (CEMOPLAF, Ecuador), Julie Forder (CARE Cambodia), Phyllis Gestrin (USAID), Evam Kofi Glover (*Planned Parenthood Association of Ghana*), Y.P. Gupta (CARE India), Lisa Howard-Grabman (*Save the Children*), Douglas Kirby (*ETR Associates*), Rekha Masilamani (*Pathfinder International*, India), Ruth Maria Medina (*Population Council*, Honduras), Dominique Meekers (*Population Services International*), Irene Moyo (JSI/SEATS), Nancy Murray (*FOCUS on Young Adults*), Mary Myaya (CARE Lesotho), Sonia Odria (*Pathfinder International*, Perú), Oladimeji Oladepo (*Department of Preventive and Social Medicine*, Nigeria), Anne Palmer (PATH Filipinas), Susan Pick de Weiss (Instituto Mexicano de Investigación de Familia y Población), Gabriela Rivera (*Pathfinder International*, México), William Sambisa (PACT Zimbabwe), Jessie Schutt-Aine (*International Planned Parenthood Federation*), Alfonso Sucrez (Fundación Mexicana para la Planeación Familiar), Oswaldo Tanaka (Organización Panamericana de la Salud), John Townsend (*Population Council/Frontiers*), Laelani L.M. Utama (*Pathfinder International*, Indonesia), Pilar Vigal (CEBRE, Chile), Amy Weissman (*Save the Children*), Anne Wilson (PATH) y Kate Winskell (*Global Dialogues*).

Los capítulos sobre la recolección de datos fueron impulsados por las presentaciones y participación del público en la reunión *YARH Measurement* patrocinada por la División de Salud Reproductiva de los *Centers for Disease Control* (CDC, Centros para el Control de las Enfermedades) y *FOCUS on Young Adults* en septiembre de 1999. Fueron especialmente útiles las presentaciones de Gary Lewis (*Johns Hopkins University Center for Communications Programs*), Paul Stupp (*CDC Division of Reproductive Health*) y Cynthia Waszak (*Family Health International*).

El equipo de Salud y Desarrollo Humano del *Education Development Center, Inc.* (EDC), se encargó del proceso de revisión bajo la supervisión experta de Phyllis Scattergood y Carmen Aldinger.

El servicio de Edición y Diseño de EDC, dirigido por Jennifer Roscoe, produjo la guía en inglés, con diseño y coordinación de Cathy Lee, y con revisiones y lectura de pruebas realizadas por el equipo editorial. Apreciamos sobremanera su creatividad.

#### Siglas y abreviaciones

|       |   |
|-------|---|
| AEC   | Área de enumeración del censo   |
| BCC   | <i>Behavior Change Communication</i> (Comunicación de cambios en el comportamiento)                                 |
| EDS   | Encuesta demográfica y de salud   |
| FLE   | <i>Family Life Education</i> (Educación en sexualidad)  |
| IEC   | Información, educación y comunicación   |
| ITR   | Infección del tracto reproductivo   |
| ITS   | Infección de transmisión sexual   |
| M&E   | Monitoreo y evaluación  |
| MT    | Medición del tamaño   |
| ONG   | Organización no gubernamental   |
| PPT   | Probabilidad en proporción al tamaño  |
| SIG   | Sistema de información gerencial  |
| SSRA  | Salud sexual y reproductiva del adolescente   |
| USAID | <i>United States Agency for International Development</i> (Agencia Estadounidense para el Desarrollo Internacional) |

*Las solicitudes para reproducir o traducir total o parcialmente esta publicación deberán dirigirse a la Unidad de Salud de Adolescentes, División de Promoción y Protección de la Salud (HPP), Programa de Familia y Población, Organización Panamericana de la Salud, 525 23<sup>rd</sup> St., N.W., Washington, DC 20037, USA.*

## Manual de monitoreo y evaluación

**Copyright © 2001 Organización Panamericana de la Salud**

**Revisión de texto: *María Virginia Pinotti, Lindsay Stewart, Francisca Infante, Dra. Matilde Maddaleno, Dra. Mariana Kastrinakis***

**Diseño: *Sylvia Singleton***

**Fotos de la portada: *Armando Waak, OPS***

# Biografías

**Susan E. Adamchak** es presidente de *Planning & Evaluation Resources, Inc.* Es experta en temas de población y desarrollo de políticas de salud, y evaluación de programas de salud reproductiva y salud pública. Recibió su Doctorado en Sociología por la Universidad de Brown.

**Katherine Bond** es Profesora Adjunta de Investigación en la Escuela de Salud Pública y Medicina Tropical de la Universidad de Tulane, y Asesora de Investigación y Evaluación en *FOCUS on Young Adults*. Ha dirigido programas de VIH/SIDA para adolescentes en EE.UU. y Tailandia, y ha capacitado a organizaciones gubernamentales y no gubernamentales en Asia y África en el uso de métodos de investigación social para diseño y evaluación de programas. Recibió su Doctorado en Salud Internacional por la Universidad de Johns Hopkins.

**Laurel MacLaren** sirvió de Coordinadora de Comunicaciones en *FOCUS on Young Adults*. Fundó y dirigió el programa de salud sexual de adolescentes en la asociación indonesia *Planned Parenthood* en Yogyakarta, y ofreció sus servicios de asistencia técnica en el desarrollo, monitoreo y evaluación de programas de salud reproductiva de adolescentes en Asia meridional. Obtuvo su Maestría en Políticas Públicas en la Universidad de Harvard.

**Robert J. Magnani**, PhD, es Profesor Adjunto del Departamento de Salud y Desarrollo Internacional de la Escuela de Salud Pública y Medicina Tropical de Tulane. Ha trabajado en el campo de población y salud internacional en las áreas de sistemas de recolección de datos y metodología, monitoreo y evaluación de programas/proyectos, y sistemas informativos de apoyo a programas de gestión y toma de decisiones sobre políticas. Ha trabajado en 27 países en vías de desarrollo en todo el mundo, con enfoque especial en el este de Asia y Asia meridional, y Latinoamérica.

**Kristin Nelson** es candidata al Doctorado en la Universidad de Tulane, y obtuvo su Maestría en Antropología Médica en la Universidad de Case Western. Ha desarrollado un trabajo minucioso en la revisión de instrumentos cualitativos y cuantitativos y programas juveniles para *FOCUS on Young Adults*. Durante dos años, vivió y trabajó en Tanzania, y ha trabajado en el área de educación para jóvenes sobre SIDA en Etiopía.

**Judith R. Seltzer** es consultora independiente y especialista en poblaciones, con énfasis en políticas demográficas, planificación familiar y salud reproductiva, y diseño y evaluación de programas internacionales de ayuda a poblaciones. Obtuvo su Doctorado en la Universidad de Johns Hopkins.

# Índice

|   |            |
|---|------------|
| <b>Prefacio</b> .....   | <b>8</b>   |
| <b>Introducción</b> .....   | <b>9</b>   |
| <i>¿Por qué monitorear y evaluar los programas juveniles?</i> .....   | <b>10</b>  |
| <i>¿Quién debería usar este manual?</i> .....   | <b>11</b>  |
| <i>Orígenes de este manual</i> .....  | <b>12</b>  |
| <i>¿Qué son el monitoreo y la evaluación?</i> .....   | <b>12</b>  |
| <br>  |            |
| <b>Sección I: Cómo hacer el monitoreo y la evaluación</b>   |            |
| <b>Preocupaciones acerca del monitoreo y evaluación de los programas de SSRA/SA</b> .....                             | <b>16</b>  |
| <i>15 desafíos del monitoreo y evaluación de los programas juveniles</i> .....  | <b>16</b>  |
| <i>Trece consejos para abordar los desafíos del monitoreo y evaluación de los programas juveniles</i> .....           | <b>23</b>  |
| <br>  |            |
| <b>Un marco para el monitoreo y evaluación de programas de SSRA/SA</b> .....  | <b>28</b>  |
| <i>Comprensión del proceso de toma de decisiones de adolescentes y jóvenes</i> .....                                  | <b>28</b>  |
| <i>Tres estrategias que promueven la salud reproductiva de los jóvenes</i> .....                                      | <b>32</b>  |
| <i>Identificación de actividades apropiadas para el programa</i> .....  | <b>40</b>  |
| <i>Aprendiendo de la experiencia internacional con la programación de la salud reproductiva de los jóvenes</i> .....  | <b>42</b>  |
| <br>  |            |
| <b>Desarrollo de un plan de monitoreo y evaluación de SSRA</b> .....  | <b>44</b>  |
| <i>Establecimiento de metas, resultados y objetivos para los programas de salud reproductiva de los jóvenes</i> ..... | <b>44</b>  |
| <i>Medición de los objetivos</i> .....  | <b>45</b>  |
| <i>Definiendo el alcance de un esfuerzo de M&amp;E</i> .....  | <b>49</b>  |
| <i>¿Qué preguntas deben hacerse para determinar el tipo de esfuerzo de M&amp;E?</i> .....                             | <b>62</b>  |
| <i>¿Qué se necesita para llevar a cabo cada tipo de evaluación? (Cómo usar el resto de este manual)</i> .....         | <b>63</b>  |
| <br>  |            |
| <b>Indicadores</b> .....  | <b>66</b>  |
| <i>¿Qué es un indicador?</i> .....  | <b>66</b>  |
| <i>Tipos de indicadores</i> .....   | <b>70</b>  |
| <i>¿Cómo deberían ser enunciados los indicadores?</i> .....   | <b>71</b>  |
| <br>  |            |
| <b>Diseños de evaluación para determinar los impactos del programa</b> .....  | <b>76</b>  |
| <i>¿Por qué se debe hacer una evaluación de impacto?</i> .....  | <b>76</b>  |
| <i>¿Cuáles son los tipos de diseños de estudio para evaluaciones de impacto?</i> .....                                | <b>77</b>  |
| <i>Experimentos al azar</i> .....   | <b>78</b>  |
| <i>Cuasiexperimentos</i> .....  | <b>82</b>  |
| <i>Diseños no experimentales</i> .....  | <b>88</b>  |
| <i>Estudios de panel</i> .....  | <b>93</b>  |
| <i>Minimizar las amenazas a la validez de la evaluación</i> .....   | <b>94</b>  |
| <i>Elección de un diseño de estudio para un programa en ejecución</i> .....   | <b>97</b>  |
| <br>  |            |
| <b>Muestreo</b> .....   | <b>102</b> |
| <i>¿Qué es el muestreo y qué papel juega en la evaluación de un programa?</i> .....                                   | <b>102</b> |
| <i>Tipos de métodos de muestreo</i> .....   | <b>104</b> |
| <i>¿Qué método de muestreo es mejor?</i> .....  | <b>104</b> |
| <i>Muestreo por conglomerados</i> .....   | <b>106</b> |
| <i>Temas clave en el muestreo por conglomerados</i> .....   | <b>110</b> |
| <i>Determinación del tamaño de la muestra</i> .....   | <b>113</b> |
| <i>Preguntas frecuentes acerca del muestreo</i> .....   | <b>118</b> |
| <br>  |            |
| <b>Plan de trabajo para el M&amp;E y la recolección de datos</b> .....  | <b>120</b> |
| <i>Preparación para la recolección de datos</i> .....   | <b>120</b> |
| <i>Métodos de recolección de datos</i> .....  | <b>126</b> |
| <i>Selección de métodos apropiados para la recolección de datos</i> .....   | <b>133</b> |
| <i>Recolección de datos</i> .....   | <b>142</b> |
| <i>Desarrollo de un plan de trabajo para el monitoreo y evaluación</i> .....  | <b>144</b> |
| <br>  |            |
| <b>Análisis de los datos de M&amp;E</b> .....   | <b>147</b> |
| <i>Procesamiento de los datos de M&amp;E</i> .....  | <b>147</b> |
| <i>Análisis de los datos de M&amp;E</i> .....   | <b>149</b> |

|   |     |
|---|-----|
| <b>Uso y diseminación de los resultados de M&amp;E</b> .....  | 169 |
| <i>Por qué usar y diseminar los resultados de M&amp;E</i> .....   | 169 |
| <i>Uso de los resultados de M&amp;E para mejorar y fortalecer el programa</i> .....                                       | 170 |
| <i>Diseminación de los resultados de M&amp;E hacia fuera</i> .....  | 171 |
| <i>¿Cómo diseminar los resultados según las necesidades de las diferentes audiencias?</i> .....                           | 172 |
| <i>Formatos comunes de diseminación</i> .....   | 173 |
| <br>  |     |
| <b>Cuadros de indicadores de SSRA</b> .....   | 177 |
| <i>¿De dónde salen los indicadores de los cuadros y cómo se pueden usar para el programa?</i> .....                       | 177 |
| <i>¿Qué tipo de indicadores se encontrarán en cada uno de los cuatro cuadros?</i> .....                                   | 177 |
| <i>¿Qué otro tipo de información se puede encontrar en los cuadros de indicadores?</i> .....                              | 179 |
| <i>Cuadro I de indicadores: Indicadores de diseño del programa</i> .....  | 181 |
| <i>Cuadro II de indicadores: Indicadores de sistemas de desarrollo y funcionamiento del programa</i> .....                | 191 |
| <i>Cuadro III de indicadores: Indicadores de implementación del programa</i> .....  | 206 |
| <i>Cuadro IV de indicadores: Indicadores de resultados de intervención del programa</i> .....                             | 218 |
| <br>  |     |
| <b>Glosario</b> .....   | 254 |
| <b>Bibliografía</b> .....   | 268 |
| <br>  |     |
| <b>Anexos</b>   |     |
| <br>  |     |
| <b>Esquemas de muestreo para estrategias de recolección de datos esenciales</b> .....                                     | 278 |
| <i>Cómo elegir una muestra sistemática de conglomerados</i> .....   | 279 |
| <i>Muestreo de conglomerados para encuestas de hogares</i> .....  | 282 |
| <i>Métodos alternativos para el muestreo de transacciones de servicios y clientes para las entrevistas de salida</i> .... | 286 |
| <i>Muestreo por conglomerados para evaluaciones/encuestas basadas en las escuelas</i> .....                               | 289 |
| <i>Cómo asignar una muestra proporcional de estudiantes a las escuelas</i> .....  | 293 |
| <i>Muestro por conglomerados para encuestas/evaluaciones de servicios de salud</i> .....                                  | 294 |
| <i>Métodos alternativos para el muestreo de transacciones de servicios y clientes para las entrevistas de salida</i> .... | 298 |
| <i>Muestreo para evaluaciones de programas de educación de pares</i> .....  | 299 |
| <i>Muestreo para encuestas de seguimiento de clientes</i> .....   | 300 |
| <i>Muestreo de grupos focales y otros esfuerzos de recolección de datos de “grupos pequeños”</i> .....                    | 302 |
| <i>Muestreo para entrevistas en profundidad</i> .....   | 302 |
| <br>  |     |
| <b>Cómo calcular los requisitos de tamaño de la muestra</b> .....   | 303 |
| <br>  |     |
| <b>Lectura recomendada</b> .....  | 307 |
| <br>  |     |
| <b>Sitios de evaluación en Internet</b> .....   | 308 |
| <br>  |     |
| <b>Sección II: Instrumentos</b>   |     |
| <i>¿Qué son las encuestas, y cómo se pueden usar para recolectar datos de M&amp;E?</i> .....                              | 311 |
| <i>¿Qué son los debates de grupos focales, y cómo pueden usarse para recolectar datos de M&amp;E?</i> ...                 | 313 |
| <i>¿Qué es una entrevista en profundidad, y cómo se puede utilizar para recolectar datos de M&amp;E?</i> ...              | 315 |
| <i>¿Qué son los clientes simulados, y cómo pueden utilizarse para recolectar datos de M&amp;E?</i> .....                  | 317 |
| <br>  |     |
| <i>1A: Lista de verificación del diseño del programa</i> .....  | 320 |
| <i>1B: Lista de verificación de la participación de los interesados</i> .....   | 322 |
| <i>1C: Lista de verificación para el curso de capacitación para el personal de proyectos de SSRA</i> .....                | 324 |
| <i>1D: Lista de verificación de las características de un servicio “amigable” hacia los jóvenes</i> .....                 | 326 |
| <i>1E: Lista de verificación de los criterios de selección</i> .....  | 327 |
| <br>  |     |
| <i>2A: Hoja mensual de recuento para consejería</i> .....   | 328 |
| <i>2B: Hoja de recuento para productos de los medios de comunicación</i> .....  | 329 |
| <i>2C: Hoja de recuento para la participación de interesados</i> .....  | 330 |
| <i>2D: Hoja de recuento de cantidad y características de los jóvenes que recibieron orientación, ene-jun 2000</i> .....   | 332 |
| <i>2E: Hoja de recuento de infraestructura institucional</i> .....  | 333 |
| <br>  |     |
| <i>3A: Formulario de informe para consejería/orientación</i> .....  | 334 |
| <i>3B: Formulario de informe para educadores pares</i> .....  | 335 |
| <br>  |     |
| <i>4: Cuestionario de movilización comunitaria (coaliciones)</i> .....  | 336 |

|  |     |
|--|-----|
| 5A: Índice de calidad de la orientación (sesiones individuales de orientación) ..... | 340 |
| 5B: Puntuación del ambiente de políticas adolescentes .....                          | 341 |
| <br>   |     |
| 6: Inventario de lugares y servicios .....   | 344 |
| Sección uno: Inventario de equipos y productos .....                                 | 346 |
| Sección dos: Condiciones del lugar de servicio .....                                 | 352 |
| Sección tres: Materiales y actividades de IEC .....                                  | 353 |
| Sección cuatro: Supervisión .....  | 354 |
| Sección cinco: Protocolos y pautas .....   | 355 |
| Sección seis: Uso de información en la administración del lugar .....                | 356 |
| Sección siete: Estadísticas de servicio .....  | 358 |
| Sección ocho: Personal .....   | 359 |
| Sección nueve: Honorarios por servicios .....  | 359 |
| <br>   |     |
| 7: Guía de observación para consejería y procedimientos clínicos .....               | 361 |
| Observación de consejería .....  | 363 |
| Métodos anticonceptivos .....  | 364 |
| Conversación sobre ITS y otros temas de salud .....                                  | 367 |
| Procedimientos médicos .....   | 368 |
| Impresiones del entrevistador sobre la consulta .....                                | 369 |
| <br>   |     |
| 8: Guía de entrevista para el personal que proporciona servicios de SSRA .....       | 371 |
| Antecedentes .....   | 371 |
| Experiencia y capacitación en SSRA .....   | 373 |
| Anticonceptivos .....  | 374 |
| Otras prácticas de salud reproductiva .....  | 378 |
| Características sociodemográficas .....  | 381 |
| <br>   |     |
| 9: Guía para las entrevistas de salida de los clientes .....                         | 383 |
| Antecedentes .....   | 383 |
| Sección uno: Características básicas .....   | 384 |
| Sección dos: Información sobre los servicios .....                                   | 385 |
| <br>   |     |
| 10: Cuestionario para interrogar a los clientes simulados .....                      | 394 |
| Antecedentes .....   | 394 |
| Preguntas para el cliente simulado .....   | 395 |
| <br>   |     |
| 11: Cuestionario para la comunidad .....   | 403 |
| Sección uno: Información sobre la comunidad .....                                    | 404 |
| Sección dos: Servicios de salud reproductiva en la comunidad .....                   | 406 |
| Sección tres: Identificación del servicio .....                                      | 408 |
| <br>   |     |
| 12: Encuesta integral de jóvenes para la guía de monitoreo y evaluación .....        | 414 |
| Introducción .....   | 415 |
| Módulo 1: Antecedentes e información relacionada .....                               | 416 |
| Módulo 2: Conocimientos sobre salud reproductiva .....                               | 419 |
| Módulo 3: ITS/VIH/SIDA .....   | 425 |
| Módulo 4: Actitudes, creencias, valores .....  | 431 |
| Módulo 5: Influencias sociales .....   | 439 |
| Módulo 6: Actividad sexual, anticoncepción y embarazo .....                          | 448 |
| Módulo 7: Habilidades y autoeficacia .....   | 461 |
| Módulo 8: Actividades de entretenimiento y preocupaciones .....                      | 465 |
| Módulo 9: Influencia de los medios de comunicación .....                             | 471 |
| Módulo 10: Drogas y alcohol .....  | 474 |
| Módulo 11: Comportamientos saludables .....  | 476 |
| Referencias de la encuesta .....   | 479 |
| <br>   |     |
| 13: Guía de debate de grupos focales con adolescentes que asisten a la escuela ..... | 480 |
| <br>   |     |
| 14: Hoja de trabajo para evaluar la efectividad de las coaliciones .....             | 482 |
| I. Estructura colaborativa y contexto comunitario .....                              | 482 |
| II. Dotación y funcionamiento del personal de apoyo .....                            | 483 |
| <br>   |     |
| 15: Cuestionario para padres de jóvenes .....  | 486 |



# □ Prefacio

**Este documento ha sido traducido y adaptado** para la Organización Panamericana de la Salud (OPS). La traducción y adaptación intentan traer al lector hispano el mensaje central del documento “A Guide to Monitoring and Evaluating Adolescent Reproductive Health Programs” producido y publicado por la organización *FOCUS on Young Adults*. Se trata de evitar, en lo posible, la traducción literal, reemplazándola por la contextual.

Las tablas y documentos de encuesta fueron traducidos más literalmente, ya que a veces, el lector tendrá que contextualizar aún más los instrumentos modelo para poder usarlos en diferentes regiones. El manual describe mecanismos recomendables para esa adaptación regional, como grupos focales para discutir el lenguaje, expresiones y “jergas” locales, y otros detalles del idioma y la comunicación.

Finalmente, aunque este es un manual producido específicamente por *FOCUS* para uso en salud reproductiva del adolescente, cabe notar que la salud “reproductiva” del joven es otra forma de referirse a la salud integral de nuestros jóvenes. Y aunque los ejemplos presentados tienden a venir del campo de la salud reproductiva, estos, en la mayoría de los casos, tienen analogías y aplicabilidad a programas de salud integral de jóvenes. Se ha indicado en la traducción hispana esa doble aplicabilidad del concepto, usando la abreviación “SSRA/SA” (salud sexual y reproductiva de adolescentes y salud del adolescente).

El lector y programador en nuestros países de Latinoamérica y Caribe puede ejercitar su juicio profesional para extender mucha de la metodología y las lecciones aprendidas de este manual a una variedad de contextos de programas que sirven a los adolescentes. La OPS los invita a poner esto en práctica y, a través de estudios de evaluación y monitoreo, a continuar la cadena de evidencia necesaria para fomentar la salud integral de nuestros adolescentes en el futuro.

**Dra. Matilde Maddaleno**, Asesora Regional  
Salud y Desarrollo de Adolescentes y Jóvenes  
**OPS/OMS**  
Diciembre 2001

# □ Introducción

En el mundo, los jóvenes están creciendo en un ambiente de cambios dinámicos. Para algunos, esta complejidad proporciona oportunidades y alternativas; para otros, significa una lucha por la supervivencia. Muchos jóvenes tienen vigor y energía, curiosidad, un sentido de aventura e invulnerabilidad. Son ingeniosos y resilientes incluso bajo las condiciones más difíciles.

Sin embargo, el período de la adolescencia es una fase en la cual los jóvenes son particularmente vulnerables a los riesgos de salud, especialmente aquellos relacionados a la sexualidad y la reproducción: VIH/SIDA, embarazos no deseados, abortos inseguros, matrimonio e hijos a muy temprana edad, infecciones de transmisión sexual y mala nutrición.

Cómo se experimenta la adolescencia y cómo afecta la salud, y en especial la salud reproductiva, está muy determinado por el momento y la secuencia en que se dan la iniciación sexual, el matrimonio y el parto, el grado de aceptabilidad o rechazo social que el momento y la secuencia de estos eventos generan, y la cantidad y disponibilidad de opciones con respecto a la educación, la capacitación en el trabajo y el empleo. En el mundo, e incluso dentro de cada país, hay una gran variedad de valores sociales y culturales que dan forma a estos eventos. Las relaciones estrechas entre los padres y la familia extendida y los jóvenes son especialmente importantes para influir en el desarrollo de los jóvenes. El acceso a servicios de salud preventivos y curativos, incluyendo la anticoncepción y el tratamiento de las infecciones de transmisión sexual, también son importantes para asegurar la salud reproductiva de los jóvenes.

## ¿Qué son los jóvenes?

Los programas que llegan a la gente joven utilizan terminologías diferentes para referirse a los jóvenes. A menudo se utiliza la palabra "adolescentes" para referirse a jóvenes entre los 10 y 19 años de edad; "jóvenes" se refiere a aquellos entre los 15 y los 24 años, y "jóvenes" a veces se usa para referirse a todos los jóvenes entre los 10 y los 24 años. Este manual los incluye a todos y usa el término "salud reproductiva de los adolescentes" (SSRA) para cubrir todos los tipos de programas.

Los programas diseñados para ayudar a los jóvenes a desarrollarse y reducir los riesgos a su salud reflejan esa variedad. Muchos de estos programas consideran a los jóvenes como un recurso crítico para el futuro, y utilizan estrategias creativas para abordar problemas complejos. Pero muchos también enfrentan financiamiento limitado, resistencia de parte de la comunidad, leyes y políticas que no brindan apoyo o falta de experiencia. Sabiendo más acerca de lo que funciona en los programas y servicios juveniles, podemos crear programas sólidos que logren sus objetivos.

La salud reproductiva se refiere a la salud y el bienestar de mujeres y hombres en términos generales y de sexualidad, embarazo, parto y condiciones y enfermedades asociadas. Muchos programas que llegan a los jóvenes están tratando de alcanzar metas de salud reproductiva que se relacionan con los efectos críticos de la salud sexual y reproductiva, tales como:

- **fecundidad:** *el número de embarazos que tiene una mujer en su vida*
- **aborto:** *en tanto se relaciona a la fecundidad y a las complicaciones de salud de las mujeres que tienen abortos inseguros o clandestinos*
- **enfermedad:** *causada por infecciones de transmisión sexual, infecciones al sistema reproductivo, VIH y condición nutricional*
- **mortalidad:** *principalmente relacionada con el embarazo y el parto, incluyendo mortalidad infantil, mortalidad materna y mortalidad relacionada al SIDA*

- **situación nutricional:** *la deficiencia nutricional afecta tanto la salud de la mujer como la de sus hijos durante la infancia.*

## □ ¿Por qué monitorear y evaluar los programas juveniles?

### ***...El monitoreo y evaluación muestran si están funcionando los programas juveniles, y cómo funcionan***

El monitoreo y evaluación (M&E) nos pueden decir si funcionan las actividades del programa, y cómo funcionan. Los administradores y donantes de los programas quieren mostrar resultados, entender cómo funcionan sus programas y evaluar cómo interactúan los programas con otros eventos y fuerzas en sus comunidades.

### ***...El M&E se puede utilizar para fortalecer programas***

Los administradores y el personal de los programas pueden evaluar la calidad de las actividades y/o servicios, y el grado en que el programa llega al público deseado. Con datos apropiados se pueden comparar localidades, fijar prioridades para la planificación estratégica, evaluar las necesidades de capacitación y supervisión, y obtener retroalimentación del público objetivo o de los participantes en el programa. Se puede dar prioridad la asignación de recursos, mejorar la información para la recaudación de fondos, proporcionar información para educar y motivar al personal, proporcionar información para la abogacía, y argumentar en favor de la efectividad del enfoque del programa.

### ***...Los resultados del M&E pueden ayudar a institucionalizar los programas***

Los resultados del M&E pueden ayudar a los interesados y a la comunidad a entender qué está haciendo el programa, cuán bien está alcanzando los objetivos, y si existen necesidades críticas que estén inhibiendo el progreso. Los resultados del M&E se pueden utilizar para educar a la junta directiva, agencias de financiamiento actuales y potenciales, funcionarios del gobierno local y miembros clave de la comunidad, tales como líderes locales, los jóvenes y sus padres que pueden ayudar a asegurar el apoyo social, financiero y político a los programas juveniles. Compartir los resultados puede ayudar al programa a establecer o fortalecer la red de individuos y organizaciones que comparten el objetivo de trabajar con los jóvenes. También puede otorgarle reconocimiento público a los interesados y voluntarios que han trabajado en pos del éxito del programa y para atraer a nuevos voluntarios.

### ***...El M&E ayuda a determinar las decisiones de las agencias de financiamiento y de las autoridades responsables***

Las agencias de financiamiento y las autoridades responsables están interesadas en el monitoreo y evaluación por varias razones. Necesitan tomar decisiones estratégicas sobre cómo gastar los recursos, y demostrar que el gasto produce resultados de calidad. Los resultados del M&E también ayudan en las decisiones acerca de la identificación y el apoyo para reproducir o expandir estrategias específicas. Los resultados del M&E a menudo revelan necesidades que no han sido satisfechas o barreras al éxito del programa, y pueden ser utilizados para abogar en favor de cambios en las políticas

o en las leyes. Los resultados del M&E pueden aumentar el reconocimiento de los programas juveniles entre el público en general y ayudar a generar percepciones positivas acerca de los jóvenes y los programas juveniles.

### ***...Los resultados del M&E contribuyen a la comprensión global de lo que funciona o da resultado***

La diseminación de los resultados del M&E—tanto los que muestran cómo está funcionando el programa como los que determinan las estrategias que no tienen el impacto deseado—contribuye a nuestra comprensión global de lo que funciona en el mejoramiento de la salud de los jóvenes. Esto promueve nuestros conocimientos en este campo creando un conjunto de lecciones aprendidas y “prácticas óptimas” que pueden fortalecer los programas de SSRA/SA alrededor del mundo.

### ***...El M&E moviliza a las comunidades en apoyo a los jóvenes***

Los resultados del monitoreo y evaluación permiten a las comunidades y a los jóvenes informar a los líderes locales sobre las necesidades de los jóvenes y abogar por financiamiento para esos programas. Los resultados obtenidos nos indican las formas en que podemos desarrollar nuevos y mejores sistemas de apoyo para los jóvenes e identificar recursos comunitarios adicionales. Además estos pueden ayudar a que la comunidad comprenda mejor los beneficios potenciales y actuales del programa y sus logros, desarrolle un sentido de propiedad a través de la participación, mejore la coordinación y movilice el apoyo a los jóvenes y al conjunto de programas que fomentan su salud y desarrollo.

## **❑ ¿Quién debería usar este manual?**

Este manual está diseñado para administradores de programas que monitorean y evalúan programas de salud de adolescentes. Algunos ejemplos de personas que pueden encontrar utilidad en este manual incluyen:

- **Administradores de proyectos a nivel comunitario:** El administrador de un programa de educación de pares en un centro comunitario juvenil puede usar este manual para establecer un sistema para monitorear la implementación de las actividades del programa.
- **Directores de programas a nivel distrital:** El director de un programa de educación de vida familiar (EVF), basado en la escuela, puede usar este manual para darle seguimiento al progreso en la implementación.
- **Administradores de salud a nivel municipal:** El administrador de un programa que busca reducir el número de embarazos e infecciones de transmisión sexual (ITS) en una clínica puede usar este manual para establecer un sistema de evaluación que

#### **Buscando ayuda externa**

La supervisión y evaluación son aspectos esenciales del desarrollo de programas de salud juvenil. Sin embargo, muchos programas no tienen la experiencia para llevar a cabo algunos aspectos de la evaluación de programas, especialmente cuando están evaluando programas grandes y complejos. Después de leer este manual, se puede buscar asistencia técnica en las universidades locales e institutos de investigación que tienen experiencia en diseño y realización de evaluaciones efectivas y eficientes.

siga los cambios en la incidencia de embarazos y de ITS entre los jóvenes de todo el municipio.

- **Funcionarios de salud a nivel estatal o provincial y administradores de organizaciones no gubernamentales (ONG):** Los funcionarios del sistema de salud a nivel estatal o provincial o administradores de ONG pueden usar este manual para recopilar información de los distritos, municipios u otras áreas geográficas o grupos de población, para tener una visión de la situación en materia de salud juvenil y de sus cambios a través del tiempo.
- **Administradores o personal técnico de voluntarios privados o agencias donantes:** Los administradores o personal técnico de voluntarios privados o agencias donantes pueden utilizar este manual para aconsejar a otras organizaciones sobre cómo mejorar sus programas, y cómo establecer un sistema de monitoreo y evaluación para programas juveniles.

## □ Orígenes de este manual

Este manual se basa en la experiencia y los conocimientos de profesionales de una variedad de disciplinas:

El campo de la planificación familiar ha sentado una sólida base para desarrollar sistemas de entrega de servicios para adultos y para medir los aportes, calidad, acceso y resultados de los programas. Este manual se apoya en las contribuciones de *The Evaluation Project* (Proyecto de Evaluación) de la Agencia de los Estados Unidos para el Desarrollo (USAID). Este proyecto aborda la evaluación centrándose en los sistemas y la entrega de un programa y en un extenso menú de indicadores de resultados en materia de salud reproductiva.

Este manual también hace uso de las lecciones aprendidas en la prevención del VIH/SIDA, con una visión abierta de la sexualidad y el comportamiento sexual, así como comprensión del valor y la contribución de las teorías de cambio social y de comportamiento en el diseño de programas efectivos para los jóvenes.

El campo del desarrollo juvenil, que ha identificado una gama de necesidades y ventajas del desarrollo, nos urge a medir las influencias sociales— más allá del conocimiento, las actitudes y prácticas individuales—tales como la creación de relaciones saludables y comunidades de apoyo, y el fomento al desarrollo de habilidades.

Asimismo, se incorporan a este manual las contribuciones del programa *FOCUS on Young Adults* de *Pathfinder International* (Focalizándose en Jóvenes) en el monitoreo de experiencias con jóvenes en el contexto de países en desarrollo. Esas experiencias han contribuido a la presentación de “elementos clave” para el diseño de programas y posibles criterios para establecer medidas de calidad y acceso a los programas.

## □ ¿Qué son el monitoreo y la evaluación?

***...El “monitoreo” y “evaluación” de procesos miden cómo está funcionando un programa***

El *monitoreo* es el seguimiento rutinario de las actividades de un programa a través de una medición regular y continua para ver si las actividades planificadas se están realizando. Los resultados

muestran si las actividades del programa se están implementado de acuerdo a los planes, y describen el grado de utilización de los servicios del programa.

La *evaluación de procesos* debería llevarse a cabo conjuntamente con el monitoreo. Las evaluaciones de procesos recogen información que mide cuán bien se están desarrollando las actividades de un programa. Generalmente esta información se recoge regularmente, por ejemplo a través de informes del personal, pero también se puede recoger periódicamente, en un esfuerzo más amplio, que incluya, por ejemplo, clientes y otros usuarios del programa, por evaluar el proceso que puede incluir el uso de grupos focales u otros métodos cualitativos. La *evaluación de procesos* se usa para medir la calidad de la ejecución de un programa y evaluar su cobertura, o el grado en que los servicios del programa están siendo utilizados por la población objetivo.

### ***...La “evaluación de resultados” y la “evaluación de impacto” miden los logros alcanzados***

La evaluación de resultados y la evaluación de impacto miden hasta qué punto se logran los resultados del programa, y evalúa el impacto del programa en la población objetivo por medio de mediciones de los cambios en el conocimiento, actitudes, comportamientos, habilidades, normas comunitarias, utilización de los servicios de salud o situación de la salud. La *evaluación de resultados* determina si los resultados en que el programa ha tratado de influir están cambiando en la población objetivo. La *evaluación del impacto* determina cuánto de los cambios observados en los resultados se deben a los esfuerzos del programa.<sup>1</sup>

Este manual tiene dos secciones, que se describen a continuación.

## **Sección I: Cómo hacer el monitoreo y la evaluación**

### **Capítulo 1: Preocupaciones acerca del monitoreo y evaluación de los programas de SSRA/SA**

*Revisa los desafíos y ofrece consejos acerca de la medición de la efectividad de los programas juveniles.*

*Trata cómo asegurarse de que los resultados sean atribuibles a los esfuerzos del programa.*

*Prevé cómo esta guía puede proveer información y apoyo al lector.*

### **Capítulo 2: Un marco para el monitoreo y evaluación de programas de la SSRA/SA**

*Revisa el estado del conocimiento sobre las influencias sociales en el comportamiento, las necesidades y cualidades del desarrollo juvenil, y los comportamientos y servicios que reflejan la búsqueda de la salud del adolescente, y resume lo que hasta la fecha sabemos que es efectivo.*

*Presenta tres estrategias que se han usado bien en la SSRA*

*Ayuda a diseñar un programa juvenil efectivo usando un “modelo lógico”.*

### **Capítulo 3: Desarrollo de un plan de monitoreo y evaluación de SSRA/SA**

*Presenta la información básica necesaria para establecer el esfuerzo de monitoreo y evaluación.*

---

<sup>1</sup> Las evaluaciones de resultados generalmente son evaluaciones de cambios en los resultados de corto y mediano plazo tales como los cambios en el conocimiento, actitudes y comportamientos. Las evaluaciones de impacto a menudo se llevan a cabo en plazos mayores y pueden identificar cambios en la salud sexual y los resultados en cuanto a la salud reproductiva en la población objetivo, tales como cambios en las tasas de enfermedades de transmisión sexual.

Proporciona pautas sobre el alcance apropiado del esfuerzo de monitoreo y evaluación y cómo planear y conducir ese esfuerzo.

| <b>¿Qué se puede determinar usando el monitoreo y evaluación?</b>  |   |
|--|---|
| <b>Monitoreo y evaluación del proceso</b>  | <b>Evaluación de resultados e impacto</b>   |
| <ul style="list-style-type: none"><li>• Si el programa está siendo implementado según lo planeado</li><li>• Calidad del programa (de acuerdo a criterios explícitos establecidos)</li><li>• Cobertura del programa</li></ul> | <ul style="list-style-type: none"><li>• Cambios en los resultados, tales como:<ul style="list-style-type: none"><li>- cambios en comportamientos</li><li>- cambio en conocimientos y actitudes</li><li>- cambios en la interacción con los padres</li><li>- cambios en las normas comunitarias</li></ul></li><li>• Si los resultados se deben a los esfuerzos del programa o a otros factores</li></ul> |

#### **Capítulo 4: Indicadores**

*Ayuda a definir y desarrollar indicadores para medir los componentes y resultados del programa.*

*Define lo que son los indicadores, proporciona ejemplos de cómo seleccionarlos y modificarlos para los objetivos y actividades del programa, y describe las fuentes de información y métodos de recolección de datos.*

*Vincula los indicadores a una muestra de instrumentos que se incluye en los anexos.*

#### **Capítulo 5: Diseños de evaluaciones para medir el impacto del programa**

*Presenta temas clave y estudios de diseños que será necesario considerar y utilizar si se está tratando de atribuir los cambios en los resultados del programa a los esfuerzos del mismo.*

*Trata la fortaleza relativa de los diseños al azar, cuasiexperimentales y no experimentales.*

*Presenta algunas opciones para comenzar una evaluación si el programa ya está en marcha.*

#### **Capítulo 6: Muestreo**

*Proporciona una visión general del tema de muestreo en la recolección de información.*

*Trata los tipos de muestreo, cómo efectuar un tipo de muestreo probabilístico especialmente útil, el muestreo por conglomerados, requisitos del tamaño de la muestra, y temas relacionados al diseño de la encuesta y el muestreo.*

#### **Capítulo 7: El plan de trabajo y de recolección de datos**

*Presenta consideraciones clave para la implementación del monitoreo y evaluación.*

*Trata la importancia de integrar el esfuerzo de monitoreo y evaluación al plan de trabajo general del programa.*

*Sugiere pasos para el desarrollo de instrumentos de recolección de datos, para probar los instrumentos y para capacitar a los entrevistadores o recolectores de información.*

*Trata consideraciones éticas.*

## **Capítulo 8: Análisis de la información de M&E**

*Trata la mecánica del análisis de información cualitativa y cuantitativa.*

*Trata el tipo de análisis que se usará para determinar y evaluar la evaluación y conclusión del proceso, resultados e impacto.*

## **Capítulo 9: Utilización y diseminación de los resultados de la evaluación**

*Presenta razones esenciales para usar y comunicar resultados del análisis de M&E.*

*Considera cómo y cuándo utilizar los resultados de la supervisión y evaluación para beneficiar a los programas.*

*Sugiere formas de comunicar los resultados.*

## **Capítulo 10: Cuadros de indicadores de SSRA/SA**

*Presenta cuatro cuadros de indicadores correspondientes a los componentes del programa: diseño, desarrollo y funcionamiento de los sistemas, implementación y resultados de la intervención.*

*Ofrece fórmulas para calcular los indicadores, posibles fuentes de información y notas que conviene tener en mente cuando se usen los indicadores.*

## **Anexos**

*Glosario*

*Bibliografía*

*Sitios relevantes en el Internet*

*Referencia de libros útiles*

## **Sección II: Instrumentos y cuestionarios**

*Proporciona instrumentos para ayudar a establecer el sistema de información.*

*Ofrece consejos para la recolección de datos usando varios métodos estándares, incluyendo encuestas, debates de grupos focales, entrevistas en profundidad y evaluación del "cliente simulado".*



# ❑ Sección I: Cómo hacer el monitoreo y la evaluación

## ❑ Capítulo 1: Preocupaciones acerca del monitoreo y evaluación de los programas de SSRA/SA

### Una mirada al capítulo:

- *Examina los desafíos y ofrece consejos sobre cómo medir la efectividad de los programas juveniles*
- *Trata acerca de cómo estar seguro que los resultados son atribuibles a los esfuerzos del programa*
- *Ofrece un avance sobre las formas en las que este manual puede ayudar*

### *...¿Cuáles son los quince desafíos del monitoreo y evaluación de los programas juveniles?*

#### 15 desafíos del monitoreo y evaluación de los programas juveniles

1. Algunos SIG no están diseñados para dar seguimiento a las características especiales de los programas juveniles.
2. Darle seguimiento a los servicios no garantiza que se sabrá a cuántos jóvenes está llegando el programa.
3. Puede existir incertidumbre acerca de si los estándares generales o las estrategias de implementación son aplicables al país en que se está trabajando.
4. Se sabe poco acerca de si los estándares para los programas de adultos son apropiados para los jóvenes.
5. Los elementos que conducen al éxito de los programas juveniles no se han documentado o diseminado apropiadamente.
6. Los programas pueden tener dificultades en desarrollar sistemas que comprendan y respondan a las necesidades de los jóvenes.
7. Medir la calidad del programa requiere comprender significados complejos y abordar asuntos delicados.
8. Medir la accesibilidad y cobertura del programa puede ser complejo.
9. Evaluar las reacciones individuales al programa puede ser difícil.
10. Medir las influencias sobre comportamientos que no ocurrieron es difícil.

11. Medir comportamientos en varios niveles de desarrollo puede ser problemático.
12. Mostrar los vínculos entre los efectos de la salud y el desarrollo juvenil puede ser complejo.
13. Algunos cambios no podrán ser medidos en mucho tiempo, y otros cambios pueden llegar a ser imposibles de medir.
14. Atribuir los cambios en los resultados a la estrategia y actividades de un programa en especial es difícil.
15. Algunos tipos de evaluaciones son costosos y pueden requerir financiamiento que va más allá de los recursos de un programa juvenil.

### **1. Algunos sistemas de información gerencial (SIG) no están diseñados para dar seguimiento a las características especiales de los programas juveniles**

Algunos SIG son parte de programas más grandes o de intervenciones de entrega de servicios y es posible que no midan las características especiales de los programas juveniles. Por ejemplo, un programa de planificación familiar que tenga un componente juvenil puede estar diseñado para darle seguimiento a la distribución de anticonceptivos; puede no estar diseñado para darle seguimiento a servicios que posiblemente sean más utilizados por los jóvenes, tales como orientación o distribución de materiales de información, educación y comunicación (IEC).

Adaptar el SIG para supervisar el programa de SSRA/SA puede requerir sólo modificaciones menores, tales como agregar una especificación de edad a los informes generados por el programa. Sin embargo, en programas de mayor escala que cubren otros grupos además de los jóvenes, agregar un nuevo componente al sistema puede ser difícil.

### **2. Darle seguimiento a los servicios no garantiza que se sabrá a cuántos jóvenes está llegando el programa**

Todos los programas necesitan determinar cómo contabilizar el número de jóvenes a los que están llegando, y cómo saber el número de jóvenes a los que se ha llegado mejorará el desempeño. Muchos programas cuentan los servicios, tales como el número de reuniones que se han efectuado o el número de preservativos que se han distribuido. Sin embargo, si todo lo que se sabe es que se han distribuido 1.000 preservativos, no se sabrá si 100 jóvenes recibieron 10 preservativos cada uno o si 500 jóvenes recibieron 2 preservativos cada uno. El sistema de seguimiento debería tratar de recoger las características clave de los participantes en el programa para ayudar a evaluar si el programa está llegando al número y tipo de jóvenes para el que fue diseñado.

Recolectar información sobre las características de la población objetivo también ayudará a entender cómo los participantes del programa cambian con el tiempo. Por ejemplo, en sus comienzos el programa puede tener como objetivo a una juventud de más edad, pero a medida que se empieza a difundir la disponibilidad de los servicios, el programa puede encontrarse sirviendo a adolescentes menores, y puede requerir ajustes en su enfoque.

### **3. Puede existir incertidumbre acerca de si los estándares generales o las estrategias de implementación son aplicables al país en que se está trabajando**

La calidad se refiere a la adecuación de un conjunto específico de actividades profesionales en relación a los objetivos que pretenden servir.<sup>2</sup> Los estándares de calidad para el diseño de programas de educación en salud se han extraído de una variedad de programas juveniles que demostraron ser efectivos para cambiar comportamientos específicos<sup>3</sup> e incluyen factores tales como:

- *un mínimo de 14 horas de instrucción,*
- *grupos pequeños y un ambiente interactivo, y*
- *modelos de práctica en comunicación, negociación y otras habilidades.*

Sin embargo, aún no sabemos hasta donde son aplicables estos estándares en un conjunto más diverso de países en desarrollo.

Las recomendaciones de este manual, tal como el Modelo Lógico que se describe en el Capítulo 2, están diseñadas para ayudar a implementar la estrategia del programa basándose en supuestos sobre qué factores sociales y de comportamiento influyen sobre los resultados de salud que se esperan obtener. Las teorías sobre las cuales se apoyan las recomendaciones están bien desarrolladas, y han pasado por un riguroso proceso para establecer cuán bien sus mediciones capturan los procesos de cambio que proponen. Aun así, la mayoría de estas teorías no se han probado en países en desarrollo y necesitan ser adaptadas a las necesidades particulares de los jóvenes en cada lugar. Debido a que las actividades del programa determinan el diseño del esfuerzo de evaluación, la falta de comprensión sobre cómo se aplican estas teorías en diferentes contextos también puede afectar la capacidad de llevar a cabo evaluaciones sólidas de resultados e impacto.

### **4. Se sabe poco acerca de si los estándares para los programas de adultos son apropiados para los jóvenes**

De forma similar, después de años de desarrollar sistemas de entrega de anticonceptivos para adultos, actualmente existen estándares de calidad más o menos aceptados. Por ejemplo, existe un consenso amplio en cuanto a que significa la entrega de servicios de anticonceptivos clínicos de calidad:

- *Competencia técnica de los proveedores de servicios,*
- *Trato respetuoso a los clientes,*
- *Comunicación efectiva con los clientes,*
- *Métodos alternativos,*
- *Mecanismos para fomentar la continuidad, y*
- *Adecuación y aceptación cultural de los servicios.<sup>4</sup>*

Sin embargo, aún no sabemos cuán universales son estos estándares para los grupos más jóvenes. Algunos de estos estándares de calidad se incluyen en los Cuadros de Indicadores como ejemplos de criterios a considerar en los indicadores de calidad, especialmente en el nivel de diseño.

---

<sup>2</sup> Green y Lewis, 1986.

<sup>3</sup> Kirby y cols., 1997.

<sup>4</sup> Bruce, 1990.

## **5. Los elementos que conducen al éxito de los programas juveniles no se han documentado debidamente**

Los programas juveniles en los países en desarrollo a menudo recurren a la intuición y la experiencia para diseñar sus programas, porque los elementos del éxito de los programas juveniles no han sido probados ni documentados. El programa *FOCUS on Young Adults* desarrolló un conjunto de documentos basados en la experiencia obtenida en los programas juveniles de países en desarrollo, los cuales apuntan a estándares que producen programas efectivos. Los siguientes “elementos clave” para el diseño de programas exitosos se identificaron en la revisión:<sup>5</sup>

- *Realizar un estudio de base para identificar temas, necesidades y público objetivo;*
- *Establecer una declaración de objetivos bien definida que contribuya al logro de los objetivos del programa; e*
- *Involucrar a los interesados locales en la planificación del programa.*

## **6. Los programas pueden tener dificultades en desarrollar sistemas que comprendan y respondan a las necesidades de los jóvenes**

Los sistemas del programa y su operación influirán sobre factores tales como el desempeño del personal, la entrega de servicios y la utilización del programa. Los sistemas del programa deben establecerse para responder a las necesidades especiales de los jóvenes. Por ejemplo, los sistemas de reclutamiento y capacitación de personal deben asegurar que el personal tenga las características y habilidades a las cuales los jóvenes responden positivamente. Los sistemas del programa ayudarán a identificar si los materiales del programa están siendo actualizados con la frecuencia necesaria para responder a los cambios en el lenguaje y las tendencias de la cultura juvenil. Un sistema de capacitación debe asegurar que los componentes necesarios para una programación juvenil están incluidos en los currículos.

## **7. Medir la calidad del programa requiere comprender significados complejos y abordar asuntos delicados**

Para determinar la calidad del programa probablemente se deban obtener interpretaciones, perspectivas y significados subjetivos de parte de los jóvenes. Estos significados son complejos porque están basados en:

- *Creencias y valores culturales,*
- *Interacciones personales dentro de una comunidad,*
- *Interacciones entre el personal del programa y su público o participantes objetivo, y*
- *Opiniones y puntos de vista de las personas que están llevando a cabo los programas.*

Por ejemplo, los programas que se centran en el empoderamiento juvenil, la movilización comunitaria, las cambiantes normas sociales y cómo influir en la cultura juvenil, necesitarán explorar los significados de temas tales como los sentimientos de autoestima, el valor del apego comunitario y la interpretación de la cultura. Puede ser difícil obtener estas reflexiones, y aún más difícil cuantificarlas. Por ejemplo, se puede contar el número de miembros de la comunidad que participa en una actividad,

---

<sup>5</sup> Birdthistle y Vince-Whitman, 1997; Israel y Nagano, 1997; Senderowitz, 1997a; Senderowitz, 1997b. Nótese que los elementos clave reflejan la experiencia de programas más enfocados a los resultados de salud reproductiva que a los resultados del desarrollo juvenil.

pero ese número puede no decir nada sobre su contribución real a la reunión, o si su preocupación aumentó como resultado de la reunión, o sobre una propuesta de estrategia para el cambio social.

Los cambios significativos en los significados y percepciones son extremadamente importantes en los programas juveniles y no deberían ser subestimados. Estos cambios juegan un papel principal en la calidad de los programas juveniles. Para capturarlos, se necesita utilizar enfoques cualitativos en la recolección de datos. Una vez que se entiendan los significados, valores y creencias relevantes, se pueden cuantificar los cambios de significados recolectando datos acerca de las variaciones en el número de participantes que comparten esas creencias

Más aun, existen numerosos obstáculos para medir los resultados de los programas de desarrollo juvenil y de salud reproductiva, lo que ayuda a explicar la razón por la cual tenemos tan poca evidencia acerca de lo que “funciona”. En primer lugar, muchos de los resultados esperados son considerados personales y privados. En algunas sociedades, hablar acerca del comportamiento sexual y las relaciones personales puede estar prohibido socialmente. En segundo lugar, los evaluadores pueden enfrentarse a la resistencia de los padres y de la comunidad al querer hacer preguntas a los jóvenes. Los líderes de la comunidad y otros interesados pueden pensar que los jóvenes en su comunidad no tienen comportamientos arriesgados; por lo tanto, sienten que no hay necesidad de hacerles preguntas. También pueden encontrar que social y políticamente es peligroso descubrir la verdad sobre el comportamiento sexual de los jóvenes, y pueden tratar de impedir la recolección de datos. Sin embargo, han habido muchos ejemplos de programas que tenían preguntas delicadas y se descubrió que los jóvenes estaban ansiosos por tratar temas de sexualidad y salud reproductiva—considerando los debates como una oportunidad para aprender y compartir sus preocupaciones y necesidades.

## **8. Medir la accesibilidad y cobertura del programa puede ser complejo**

El acceso a los programas de salud reproductiva se refiere al grado en que los jóvenes pueden obtener servicios apropiados de salud reproductiva, con un nivel de esfuerzo y costo que es aceptable y que está dentro de las posibilidades de la gran mayoría de los jóvenes en una población dada.<sup>6</sup> Podemos definir *accesibilidad* de varias maneras:

- **Geográfica/física:** *Horario conveniente, ubicación, amplia gama de servicios necesarios.*
- **Económica:** *Precios razonables.*
- **Psicosocial:** *Percepción de privacidad, percepción de que los hombres y mujeres jóvenes, casados o solteros son bienvenidos, sensación de seguridad y confidencialidad, percepción de que los proveedores están interesados, informados y son sensibles a las necesidades de los jóvenes.*
- **Administrativa:** *El programa tiene personal especialmente capacitado, respeto a los jóvenes, tiempo adecuado para la interacción, participación de los jóvenes en el diseño y en la continua retroalimentación y tiempos de espera cortos.*

La *cobertura* se refiere al grado en que los servicios del programa—educacionales o clínicos—están siendo utilizados por la población objetivo. La cobertura se puede medir de las siguientes maneras:

- *determinando la proporción de la población objetivo a la que el programa está llegando, o*
- *determinando las características de la población a la que el programa está llegando.*

---

<sup>6</sup> Bertrand et al., 1995.

Algunos aspectos de la accesibilidad y cobertura se pueden medir por la ausencia o presencia de algo, y medirlo puede ser relativamente fácil. Por ejemplo, averiguar si el programa tiene un horario conveniente y tarifas razonables puede ser fácil de determinar con una pequeña encuesta de la población objetivo. Sin embargo, medir asuntos más subjetivos que involucran juicios—como por ejemplo si el personal respeta a los jóvenes—puede ser más difícil, porque muchos jóvenes pueden dudar de dar su opinión verdadera acerca del personal del programa por temor a consecuencias negativas, tales como ser excluidos de los servicios.

De igual forma, determinar algunas características de los jóvenes puede ser sencillo, como por ejemplo preguntarle a los participantes sobre su edad, sexo y lugar de residencia. Sin embargo, si el programa está llegando a grupos específicos de jóvenes, especialmente los marginados, puede ser más difícil recolectar esa información. Por ejemplo, si el programa pretende llegar a jóvenes que han sido sexualmente abusados, el tema puede ser demasiado sensible como para que los participantes respondan. Quizás se tendrá que preguntar repetidamente, y asegurarles a los participantes que no corren riesgo alguno por dar su opinión.

### **9. Evaluar reacciones individuales al programa puede ser difícil**

Una medida de calidad es cómo está siendo recibido el programa por los interesados, el personal y los participantes. Evaluar cómo está siendo recibido el programa por estos grupos contribuirá a encontrar formas de superar la resistencia social a los programas juveniles. También ayudará a determinar si el programa va por buen camino, y a identificar problemas a tiempo para poder corregirlos. Sin embargo, es difícil obtener y analizar reacciones individuales al programa.

Por ejemplo, se podría involucrar a los jóvenes y a los miembros de la comunidad para que piensen de forma crítica acerca de sus necesidades, y para que consideren cómo podría el programa llegarles de mejor manera. Empero, algunos individuos pueden tener problemas para articular sus necesidades, o sus opiniones pueden oponerse a lo que sabemos sobre los factores que afectan los resultados de salud. Por ejemplo, algunos miembros de la comunidad pueden pensar que es peligroso entregar información sobre salud reproductiva a los jóvenes, y pueden querer censurar los medios de comunicación para generar resultados de salud positivos entre los jóvenes. Otros pueden expresar automáticamente puntos de vista que están en línea con las normas y valores sociales, aun cuando no reflejen las verdaderas necesidades de la comunidad. Los jóvenes pueden ser particularmente renuentes a expresar reacciones negativas acerca del programa a los evaluadores, que a menudo son mayores y tienen más autoridad.

Se pueden encontrar tendencias similares en las reacciones del personal y los voluntarios del programa. Los procesos de evaluación estimulan al personal a reflexionar sobre su trabajo, ver las fortalezas y debilidades del programa y considerar estrategias alternativas. Aun así, aunque la mayor parte de las personas que trabajan con jóvenes están profundamente involucradas y comprometidas, algunos tienen un enfoque más ideológico. Pueden pensar que sus estrategias están funcionando aunque exista poca evidencia para sugerir que esto ocurre. Por ejemplo, parte del personal puede insistir en que aumentar el acceso a los servicios de anticonceptivos es la mejor manera de producir resultados, ignorando el hecho que para los jóvenes que se abstienen de tener relaciones sexuales un servicio más importante podría ser obtener apoyo para sus reflexiones y su decisión de abstenerse del sexo. Otros pueden sentir que su compromiso y su duro trabajo debería traducirse en resultados, y encuentran desmoralizante hablar sobre cómo sus esfuerzos pueden estar mal encaminados. El personal requerirá de un ambiente confiable y un proceso de apoyo para permitir el tipo de reflexión que lleve a admitir que las estrategias del programa pueden necesitar modificaciones.

### **10. Medir las influencias sobre comportamientos que no ocurrieron es difícil**

Muchos programas de SSRA/SA se enfocan en la prevención de comportamientos poco sanos y de influir sobre el camino al desarrollo. A menudo se preocupan de medir los eventos que no ocurrieron por la intervención del programa. Por ejemplo, algunos programas pueden tener como objetivo retrasar

el comienzo de la actividad sexual o prevenir el sexo no deseado. Otros pueden tratar de prevenir los matrimonios precoces, demorando así la primera experiencia sexual de las mujeres jóvenes, y aumentando la edad del primer nacimiento a un momento en que el parto sea más seguro. Obviamente, medir la ausencia de ciertos comportamientos es complejo. Requiere estimar qué nivel de determinados comportamientos habrían existido si no hubiera habido una intervención, para así explicar porqué una intervención evitó un determinado comportamiento.

### **11. Medir comportamientos en varios niveles de desarrollo puede ser problemático**

A pesar de que los programas juveniles se centran en llegar a los jóvenes a lo largo de la transición hacia el desarrollo, no siempre se sabe qué resultados deberían esperarse a edades específicas. Por ejemplo, es posible no saber cuál es la edad promedio de la primera experiencia sexual en la población objetivo. Sin embargo, medir los resultados sobre el comportamiento sexual puede ser problemático. Algunos jóvenes pueden no saber acerca de ciertos comportamientos sexuales y, por lo tanto, pueden tener problemas para contestar preguntas acerca de ellos. Esto podría sesgar los resultados (por ejemplo, cuando una muchacha que se ha tomado de la mano con un muchacho informa que se ha involucrado en "actividad sexual"). Los miembros de la comunidad y a veces el mismo personal del programa pueden sentir que no es apropiado presentar nuevos temas a los jóvenes, tales como el comportamiento sexual o comportamientos ilegales, a través del esfuerzo de recolección de datos.

### **12. Mostrar el vínculo entre los efectos de la salud y el desarrollo juvenil puede ser complejo**

Muchos programas están cada vez más enfocados a vincular los resultados de salud con el desarrollo juvenil. Por ejemplo, un programa puede querer demostrar que un aumento en la educación de las niñas puede ayudar a demorar la primera experiencia sexual y, por lo tanto, tiene un resultado de salud positivo. Sin embargo, qué aspectos del desarrollo juvenil influyen sobre los resultados de salud puede ser difícil de predecir. No se puede deducir que los factores del desarrollo tendrían la misma influencia en la salud en diferentes lugares, ya que los resultados están integrados en contextos locales específicos, cada uno con sus valores sociales y culturales propios. Medir el contexto social y cultural del desarrollo juvenil es difícil y puede requerir tiempo y recursos que muchos programas no tienen.

### **13. Algunos cambios no se podrán medir en mucho tiempo, y otros pueden no llegar a medirse nunca**

Algunos cambios pueden tardar mucho tiempo en aparecer. Por ejemplo, pueden pasar varios años antes de que se puedan observar cambios en la situación de salud de los jóvenes, en comparación con el relativamente poco tiempo que toma observar cambios, tales como cambios en el nivel de conocimiento. Más aún, algunos cambios en los resultados pueden ocurrir mucho tiempo después de que el programa haya terminado; por ejemplo, es posible que un programa que promueve el retraso de la primera relación sexual entre los jóvenes de 10 a 12 años no observe sus resultados hasta varios años después de que los participantes tomaran parte en el programa, a la edad en la que se puede esperar el comienzo de la actividad sexual. Por lo tanto, es importante darle seguimiento a las tendencias de estos comportamientos. Para muchos resultados de gran interés, no se sabe cuánto tiempo tomará producir cambios. Aun así, se espera que muchos programas juveniles demuestren cambios en los resultados de largo plazo en breves períodos de tiempo. Algunos programas definen sus objetivos de forma poco realista y después concluyen equivocadamente que los programas no tuvieron éxito, cuando en realidad se necesita más tiempo para demostrar los cambios.

De igual forma, algunas estrategias de programas, particularmente aquellas que abordan el cambio social, son más difíciles de medir en términos numéricos o cuantificables. Por ejemplo, medir procesos sociales complejos, tales como la movilización y el empoderamiento de la comunidad, puede

ser difícil debido a que conceptualmente no estamos seguros de cómo definir estos procesos ni de expresar cómo están sucediendo.

#### **14. Atribuir los cambios en los resultados a la estrategia y actividades de un programa en especial es difícil**

¿Cómo se puede llegar a la conclusión de que los cambios observados en la población objetivo obedecen a las actividades del programa? Medir solamente los cambios en los resultados no es suficiente para concluir que los cambios son el resultado del programa. Por ejemplo, otros eventos, tales como cambios de las condiciones económicas o sociales, podrían tener un impacto en los resultados que se están tratando de medir. También pueden haber otras actividades del programa dirigidas a la audiencia objetivo, tales como campañas de los medios de comunicación, que se están realizando al mismo tiempo. Finalmente, el programa pudo haber atraído participantes que estuvieran predispuestos a los resultados positivos que el programa estaba fomentando. La principal manera de determinar que un cambio observado en los indicadores de resultados se puede atribuir al programa es usando un sólido diseño de estudio (ver Capítulo 5). Aun así, diseñar e implementar un sólido diseño de estudio requerirá un alto nivel de recursos y habilidades, y puede no ser factible para algunos programas.

#### **15. Algunos tipos de evaluaciones son costosas y pueden requerir financiamiento que va más allá de los recursos de un programa**

Las evaluaciones de resultados y de impacto pueden ser costosas, especialmente cuando miden varios resultados o aquellos que son difíciles de evaluar. Si los programas no pueden depender de las fuentes de datos existentes, es posible que deban recolectar mucha información nueva acerca de la población joven a la cual quieren llegar. Desarrollar instrumentos de estudio, llevar a cabo un muestreo correcto y recolectar datos de los individuos puede ser costoso. Los programas que no tengan capacidad interna para efectuar evaluaciones también pueden enfrentar el costo adicional de asistencia técnica o de contratar evaluadores externos.

### ***...Trece consejos para abordar los desafíos del monitoreo y evaluación de los programas juveniles***

El personal y los evaluadores de programas alrededor del mundo están mejorando sus habilidades y desarrollando soluciones creativas a los duros desafíos del monitoreo y evaluación. A continuación se presentan consejos de los profesionales que hacen trabajo en el terreno y se ofrecen pautas sobre cómo usar este manual para abordar los desafíos de las mediciones.

#### **13 consejos que abordan los desafíos del monitoreo y evaluación de los programas juveniles**

1. Monitorear si el sistema está entregando lo que debe: programas para los jóvenes.
2. Basar las actividades del programa y, por ende, la evaluación en la teoría.
3. Revisar lo que se sabe acerca de los factores que influyen sobre los resultados de la salud.
4. Probar y documentar los elementos que contribuyen a la efectividad del programa.
5. Comprometerse a un proceso genuinamente participativo.
6. Asegurar que la recolección de datos responde a preocupaciones éticas.



7. Ser creativo al hacer preguntas delicadas.
8. Definir los objetivos de forma realista y darse tiempo suficiente para medir los cambios.
9. Usar un enfoque cualitativo-cuantitativo combinado.
10. Usar los datos generados por el proceso de monitoreo y evaluación para sustentar los hallazgos de la evaluación de resultados e impactos.
11. Aprender por el método de ensayo y error.
12. Limitar los costos de la evaluación cuando sea posible.
13. Basarse en las ventajas al evaluar los programas juveniles.

### **1. Monitorear si el sistema está entregando lo que debe: programas para los jóvenes**

Monitorear las partes del sistema del programa que responden específicamente a las necesidades de los jóvenes. En los cuadros de indicadores del Capítulo 10, se proporcionan algunas notas sobre cómo considerar temas en el desarrollo de un nuevo sistema o para adaptar un sistema existente que pueda captar las necesidades de los programas juveniles.

### **2. Basar las actividades del programa y, por ende, la evaluación en la teoría**

Basar las estrategias del programa en la teoría ayuda a explicar cómo están funcionando los programas y, si estos son exitosos, ayuda a repetirlos y adaptarlos. El Modelo Lógico, que se introduce en el Capítulo 2, es un ejemplo de cómo un programa puede planificar sus actividades basándose en las teorías del comportamiento en la salud y de cambio social. Una mejor comprensión sobre cómo estas y otras teorías se pueden aplicar en distintos contextos fortalecerá la capacidad de llevar a cabo evaluaciones de resultados e impactos científicamente sólidas.

### **3. Revisar lo que se sabe acerca de los factores que influyen sobre los resultados de la salud**

Para ayudar a demostrar el vínculo entre resultados de salud y necesidades del desarrollo, se debe revisar, en primer lugar, lo que se sabe acerca de las influencias que se asume afectan los resultados. Como lo sugiere el Capítulo 2, la mejor forma de hacerlo es revisando la literatura e investigaciones existentes acerca de la población objetivo. Sin embargo, si no se puede acceder a la literatura publicada, o no está bien desarrollada en la localidad, se pueden revisar las experiencias del personal o conversar con colegas de otras organizaciones. También se pueden encontrar formas de evaluar estas influencias a través de una recolección de datos creativa, tal como hacer preguntas sobre un comportamiento específico de varias maneras distintas, o modificando el lenguaje y la terminología en los instrumentos de encuesta, para reflejar los temas que son importantes en la localidad.

### **4. Probar y documentar los elementos que contribuyen a la efectividad del programa**

Varios elementos contribuyen a la efectividad de un programa juvenil. Por ejemplo, los elementos de diseño propuestos en los cuadros de indicadores reflejan el estado actual del conocimiento acerca de cuáles son las características del diseño clave para el éxito de un programa. Se sugiere que estos elementos sean probados o modificados de acuerdo al lugar específico y las prioridades del

programa. Probar significa que los elementos del diseño deberían ser usados y evaluados para ver si son apropiados y efectivos.

Los sistemas que se establezcan para implementar un programa juvenil también son claves para el éxito del programa, y pueden ser más complejos que los sistemas utilizados para implementar otros programas de servicios de salud reproductiva. Más aún, los criterios para evaluar la calidad del funcionamiento de un programa juvenil no han sido probados sistemáticamente. Los posibles criterios para evaluar la calidad del sistema y su operación se sugieren en los cuadros de indicadores, e incluyen elementos tales como:

- *Reclutamiento de personal con las habilidades apropiadas*
- *Componentes del programa de capacitación*
- *Participantes del programa de capacitación que han dominado las habilidades*
- *Contenido del programa de salud reproductiva*
- *Desempeño del personal*

La mejor forma de establecer programas y servicios de calidad para los jóvenes es incorporando los elementos identificados en la experiencia internacional de los programas juveniles, las lecciones del campo de la planificación familiar y la intuición y experiencia propias.

### **5. Comprometerse a un proceso genuinamente participativo**

Las evaluaciones que comprometen e involucran a los interesados y al personal con toda probabilidad producen reacciones que son más críticas y honestas que aquellas llevadas a cabo exclusivamente por expertos externos. Un proceso participativo también motiva a la comunidad y al personal a utilizar la información del proceso de evaluaciones, y a tener un sentido de propiedad sobre los resultados de las evaluaciones. Darle a los jóvenes y adultos la oportunidad de debatir y analizar sus preocupaciones, y sugerir e implementar soluciones, también puede aumentar la efectividad del programa para alcanzar sus objetivos.<sup>7</sup> En este manual se proporcionan algunos consejos sobre cómo comprometer a los jóvenes, a los miembros de la comunidad y a otros interesados.

### **6. Asegurar que la recolección de datos responde a preocupaciones éticas**

En la conducción de investigaciones y estudios de evaluación se deben ejercitar estándares profesionales de conducta, así como principios y valores morales. Las revisiones éticas están diseñadas para considerar y reducir los riesgos potenciales y las consecuencias negativas para los involucrados, como resultado de su participación en un estudio o evaluación. Responder a las preocupaciones éticas mejorará la relación con la comunidad y aumentará la capacidad de recolectar información de calidad. Mientras más ético sea el esfuerzo de recolección de información, más honesta y confiable será la información recolectada, asegurando que los resultados de M&E sean válidos. En el Capítulo 7 se tratan formas éticas de recolectar información de los jóvenes.

### **7. Ser creativo al hacer preguntas delicadas**

Aunque es difícil hacer preguntas de una naturaleza sensible, se pueden hacer exitosamente en muchos ambientes diferentes. En primer lugar, puede necesitarse el apoyo de una amplia gama de organizaciones comunitarias a las cuales será necesario explicar claramente porqué es necesario hacer

---

<sup>7</sup> La literatura acerca de procesos participativos de evaluaciones está bien desarrollada. Un recurso específico al contexto de los programas de salud reproductiva de los jóvenes es Shah et al., 1999. *Listening to Young Voices: Facilitating Participatory Appraisals on Reproductive Health with Youth*. (Escuchando las Voces Jóvenes: Facilitando las Evaluaciones Participativas en la Salud Reproductiva de los Jóvenes). Washington DC: *FOCUS on Young Adults*.

las preguntas. En segundo lugar, se necesitará el consentimiento de los padres, particularmente para los jóvenes que legalmente son menores de edad. En tercer lugar, se pueden usar “patrones de omisión” para evitar preguntas explícitas de contenido sexual sobre uso de anticonceptivos a jóvenes que no han tenido relaciones sexuales. Se proporcionarán consejos adicionales en la discusión sobre recolección de datos en el Capítulo 7.

### **8. Definir los objetivos de forma realista y darse tiempo suficiente para medir los cambios**

Los programas que definen sus objetivos en forma poco realista pueden llevar a falsas conclusiones. Debe presupuestarse suficiente tiempo antes de tratar de medir los cambios en resultados, y asegurar que los objetivos claramente expliquen los resultados que se intentarán lograr.

### **9. Usar un enfoque cualitativo-cuantitativo combinado**

Los métodos cualitativos se pueden utilizar para definir los contextos sociales y culturales, y para desarrollar vocabularios para los programas de educación en salud, cada uno de los cuales contribuye a la formulación de instrumentos a ser utilizados durante las encuestas cuantitativas. Los métodos cuantitativos aseguran una recolección de datos estandarizada a lo largo del tiempo, y permiten aplicar medidas definitivas del cambio en los resultados que se pueden generalizar a una mayor población. También se pueden utilizar para mostrar que los cambios se deben a las actividades del programa. Entonces se pueden usar los datos cualitativos para explicar los hallazgos de las encuestas cuantitativas, y pueden revelar resultados del programa que no fueron detectados a través de los métodos cuantitativos.

Los métodos cualitativos también se pueden utilizar para evaluar las metas del programa que son difíciles de cuantificar, tales como el empoderamiento y el cambio social. Por ejemplo, un enfoque cualitativo es una “historia natural”, un método que solicita a los participantes y al personal que describan la evolución del programa. Utilizar este método puede ayudar a comprender qué cambios se lograron y por qué. Los resultados de la historia natural se pueden utilizar para desarrollar un enfoque cuantitativo que mida si los cambios están produciendo el resultado esperado en la comunidad más amplia.

En el Capítulo 5 sugerimos utilizar una combinación de enfoques cualitativos y cuantitativos para desarrollar indicadores y recolectar datos, los cuales ayudarán a abordar algunas de estas preocupaciones en más detalle.

### **10. Usar los datos generados por el proceso de monitoreo y evaluación para sustentar los hallazgos de la evaluación de resultados e impactos**

Efectuar evaluaciones de resultados e impactos requiere de recursos y tiempo, e incluso aquellas que están bien diseñadas pueden mostrar resultados que no son concluyentes. Utilizar los datos del monitoreo y evaluación de procesos puede fortalecer los resultados de las evaluaciones de resultados e impactos. Por ejemplo, puede que el SIG recoja información acerca de la exposición a los servicios del programa, tal como el contacto con adolescentes educadores. Si se está tratando de demostrar resultados tales como un aumento en el uso de preservativos, se puede medir si los jóvenes recibieron preservativos o si fueron referidos por un adolescente educador. Aunque esto no dará evidencias concluyentes acerca del efecto del programa sobre el total de la población objetivo, puede ayudar a demostrar asociaciones entre los resultados positivos y la exposición a las actividades del programa.

La información que se recoja a través del monitoreo y evaluación del proceso también ayudará a fortalecer los argumentos que defienden que los cambios fueron resultado del programa, incluso si una evaluación de impacto no es factible. Por ejemplo, el SIG puede mostrar que ciertas actividades fueron ejecutadas con más frecuencia que otras. La evaluación de procesos puede determinar que a los

jóvenes les gustaron más algunos mensajes que otros, y como resultado, se involucraron más en el programa. También puede que el SIG documente que el apoyo de los líderes de la comunidad a actividades específicas dio como resultado un aumento en la participación o la creación de nuevas actividades. Documentar factores tales como utilización de servicios, participación en el programa y reacción a las estrategias del programa, fortalecerá el argumento que expone que el programa produce resultados.

### **11. Aprender por el método de ensayo y error**

Para otros asuntos relacionados con las mediciones, se necesita aprender con el método de ensayo y error. Por ejemplo, cuando nos damos cuenta de que se pueden hacer preguntas acerca del comportamiento sexual, aun en lugares con valores muy tradicionales. En cada lugar puede ser necesario modificar quién hace las preguntas, de qué forma hace las preguntas o el lugar en que se hacen las preguntas. Se debe ser creativo al generar autoinformes válidos sobre comportamientos arriesgados. A menudo se obtiene una respuesta “socialmente aceptada” en vez de un informe veraz sobre un comportamiento.

En esta etapa, muchas de las medidas sugeridas en este manual aún no han sido probadas. Hay que basarse en lo que ya se sabe a medida que se recopile más evidencia y a medida que los programas intenten enfoques más sistemáticos de monitoreo y evaluación.

### **12. Limitar los costos de la evaluación cuando sea posible**

Aunque los costos de la evaluación de resultados e impactos pueden ser altos, hay formas en que los programas pueden limitar los costos y producir resultados válidos. Por ejemplo, una evaluación puede examinar sólo los resultados más importantes para el programa. También se pueden reducir los costos midiendo los resultados que requieren métodos de recolección de datos menos costosos o utilizando datos ya existentes. Para algunos programas, puede ser factible capacitar y utilizar al personal para llevar a cabo algunas partes de la evaluación. Una estrategia sólida de muestreo puede ayudar a limitar la cantidad de datos recolectados sin comprometer la validez de los resultados de la evaluación.

### **13. Basarse en las ventajas al evaluar los programas juveniles**

A pesar de que los desafíos descritos en este capítulo son muchos, la tarea de evaluar los programas de salud reproductiva de los jóvenes puede ser muy gratificante. Asimismo, pueden haber otras formas relativamente simples para evitar los desafíos en la medición de resultados, incluyendo la asignación al azar de jóvenes en escuelas, ya sea individualmente o por salas de clase, y darles seguimiento. También se pueden encontrar comunidades donde la demanda del programa sea mayor que la oferta.

Se puede elegir evaluar el impacto de programas seleccionados, no todos los programas, o usar un diseño de evaluación de la intervención después del tratamiento, que se describe en el Capítulo 5. Si unos pocos jóvenes necesitan el programa desesperadamente, se les puede permitir estar en el programa pero no en el estudio. Finalmente, hay muchas ventajas de trabajar con jóvenes. Están interesados en aprender y los cambios se pueden producir relativamente rápido. A pesar de que pueden estar en un período de mayor vulnerabilidad, mejorar los resultados en los jóvenes es una inversión en nuestro futuro.

## □ Capítulo 2: Un marco para el monitoreo y evaluación de programas de SSRA/SA

### Una mirada al capítulo:

- *Considera los múltiples factores que le dan forma a la adolescencia*
- *Presenta tres estrategias importantes utilizadas para mejorar la salud reproductiva de los jóvenes*
- *Trata el Modelo Lógico, un enfoque para diseñar una estrategia efectiva*

### *...Comprensión del proceso de toma de decisiones de adolescentes y jóvenes*

#### *La adolescencia no es igual en todas partes*

La definición de adolescencia—e incluso su existencia—ha sido sujeto de debate por mucho tiempo. Algunos argumentan que la adolescencia es un período en el que los niños alcanzan la madurez física pero no están agobiados con los roles y responsabilidades de los adultos. Ellos dicen que la adolescencia es un fenómeno de las sociedades industriales modernas.<sup>8</sup> Otros han teorizado que la adolescencia existe en todas las culturas y en todos los tiempos, y definen la adolescencia como una fase de la vida que involucra el manejo de la sexualidad entre individuos solteros, la organización social y la influencia de los grupos de pares entre los adolescentes, y la capacidad de enfrentar la vida.<sup>9</sup> Una modificación reciente de esta última definición dice que la adolescencia es una época de mayor vulnerabilidad para las muchachas y de formación de la capacidad crítica para los jóvenes de ambos sexos (10 a 19 años), sin importar su estado civil y/o maternidad.<sup>10</sup>

La adolescencia se experimenta de distintas maneras en cada sociedad, e incluso dentro de las sociedades pueden existir grandes diferencias en la forma en que algunos jóvenes experimentan la adolescencia comparados con otros. Para desarrollar los resultados, objetivos e intervenciones de los programas para que estos tengan el impacto deseado, primero debe entenderse el contexto específico de la población juvenil objetivo con la que se planea trabajar.

Los factores socioculturales influyen en cómo los jóvenes experimentan la adolescencia, y el comportamiento sexual de los adolescentes refleja la variedad de normas y expectativas. Particularmente en los lugares en donde se han producido cambios sociales considerables en las últimas décadas, los jóvenes luchan por equilibrar los mensajes mixtos y tratan de encontrar lo que es mejor para ellos.

#### *Una amplia gama de factores influye en la salud reproductiva de los jóvenes*

Los factores que influyen en cómo los jóvenes experimentan la adolescencia abarcan cinco grandes categorías:

- *las características individuales de los jóvenes, incluyendo sus conocimientos, actitudes, creencias, valores, motivaciones y experiencias*

---

<sup>8</sup> Caldwell, 1998.

<sup>9</sup> Schlegel, 1995.

<sup>10</sup> Mensch y cols., 1998a: 77.

- *parejas sexuales y pares*
- *familias y adultos en la comunidad*
- *instituciones que apoyan a los jóvenes y les proporcionan oportunidades, tales como escuelas, lugares de trabajo y organizaciones religiosas*
- *comunidades a través de las cuales se transmiten las expectativas sociales sobre normas de género, comportamiento sexual, matrimonio y maternidad*

## Hallazgos de las investigaciones

Los resultados de investigaciones presentados aquí representan una síntesis de más de 350 estudios; alrededor de 250 se efectuaron en los Estados Unidos y cerca de 100 en Asia, África y América Latina y el Caribe. Cada investigación que se terminó después de 1975 tuvo una muestra de más de 100 jóvenes, se usaron criterios científicos y se revisaron los antecedentes de la edad en la primera relación sexual, frecuencia de la actividad sexual, número de parejas sexuales y uso de preservativos y anticonceptivos. La investigación identificó factores protectores y factores de riesgo. Los estudios de los Estados Unidos fueron sintetizados por Douglas Kirby, de *ETR Associates* (Kirby, 1999b), y la mayoría de los estudios de los países en desarrollo fueron revisados por Ilene Speizer y Stephanie Mullen, de la Universidad de Tulane (Speizer y Mullen, 1997). Hay resultados adicionales que provienen de trabajos por publicar de FOCUS, incluyendo "Influencias Sociales sobre los Comportamientos Sexuales de los Jóvenes en Lusaka, Zambia", "Factores Protectores en Contra de los Comportamientos Sexuales Riesgosos entre los Estudiantes Secundarios Urbanos en Perú" y "Las Influencias de los Contextos Familiares y de Pares sobre los Comportamientos Sexuales de los Jóvenes Solteros en Ghana".

Estos factores influyen en cuánta educación debería recibir un joven, cuáles son los patrones de noviazgo y matrimonio, y cuándo se supone que una persona debe adquirir las responsabilidades de un adulto, tales como trabajar y mantener la familia. Sin embargo, a menudo también están en conflicto entre ellos. Por ejemplo, las normas de los pares acerca de cuán apropiadas son las relaciones entre muchachos y muchachas pueden ser muy distintas a las de la familia y la comunidad. Más aún, cada uno de estos factores está cambiando constantemente a medida que cambia el mundo. Comprender y responder a estos factores es una parte importante del desarrollo de programas efectivos de SSRA/SA.

### **La investigación revela mucho acerca de cómo estos factores dan forma a los procesos de toma de decisiones reproductivas entre los adolescentes**

Los investigadores están prestándole cada vez más atención a los *antecedentes*, factores que preceden e influyen en cómo los adolescentes toman decisiones respecto a los comportamientos sexuales y de salud. Los antecedentes pueden ser positivos, "factor protector", o negativos, "factor de riesgo". Si bien la investigación puede mostrar la relación de los antecedentes con la toma de decisiones sexuales, es más difícil identificar cuáles son los antecedentes que más influyen en los resultados de salud reproductiva.

A continuación se discuten los resultados de investigación en cada uno de los cinco ámbitos de influencia.

## Características individuales

En algunos casos, los jóvenes pueden calcular o negociar los riesgos antes de tomarlos. Pueden decidir tomar los riesgos porque se sienten invulnerables, no saben las consecuencias, quieren experimentar, o porque involucrarse en riesgos les da posición social o beneficios monetarios.

### **Riesgos sexuales**

Los riesgos sexuales son comportamientos sexuales que ponen a un individuo en riesgo de embarazo no planificado, ITS/VIH o problemas de salud relacionados al embarazo y la maternidad. Los riesgos sexuales específicos incluyen:

- *Inicio precoz de la actividad sexual*
- *Relación sexual sin uso de anticonceptivos*
- *Relación sexual sin uso de preservativos*
- *Relación sexual con más de una pareja*
- *Relación sexual con una pareja infectada por ITS o VIH*

La investigación ha encontrado que el nivel de conocimientos sobre salud sexual y reproductiva, y las normas y valores de la comunidad y la familia acerca de la salud sexual y reproductiva afectan las decisiones de salud reproductiva de los adolescentes. Por ejemplo, las mujeres jóvenes en Ghana le dan un gran valor a la fertilidad precoz, lo cual es un factor de riesgo para los embarazos precoces. La autoeficacia, el desempeño académico y la motivación por ser exitoso en la escuela parecen proteger a los jóvenes de los riesgos sexuales. Los jóvenes que están activamente involucrados en el aprendizaje, que le dan un alto valor a ayudar a las personas, y que aceptan responsabilidades y se hacen responsables tienen menos probabilidades de incurrir en riesgos sexuales. Las intenciones del comportamiento a menudo afectan la toma de riesgos por parte de los adolescentes; por ejemplo, es menos probable que los jóvenes que quieren evitar las infecciones de transmisión sexual (ITS) incurran en riesgos sexuales.

Se han relacionado otros riesgos a los comportamientos sexuales arriesgados entre los jóvenes. El uso de alcohol y drogas, fumar, depresión y estrés, la soledad y abandonar el hogar aumentan la toma de riesgos sexuales entre los jóvenes. A menudo es más probable que los jóvenes que han sido víctimas de abuso sexual o físico durante la niñez o adolescencia tomen riesgos sexuales.

Existen factores biológicos que también parecen contribuir a la toma de riesgos de los adolescentes. El desarrollo físico precoz y los altos niveles de testosterona aumentan la toma de riesgos. La edad y el género también influyen sobre la toma de riesgos sexuales; en general, es más probable que los hombres jóvenes tomen más riesgos sexuales que las mujeres jóvenes, y lo mismo ocurre con los jóvenes de más edad.

## Pares y parejas sexuales

Los investigadores han encontrado que, si los jóvenes piensan que sus amigos tienen relaciones sexuales, consumen alcohol o drogas, o fuman, es más probable que ellos adopten esos comportamientos. Los desequilibrios de poder en una relación, tales como edad y diferencia de ingresos entre la pareja, intercambio de dinero o de otros bienes a cambio de sexo y la presión sexual de una pareja también contribuyen al aumento de los riesgos sexuales. A la inversa, un sentido de compromiso en una relación parece proteger a los jóvenes de resultados de salud indeseados. Existe evidencia que sugiere que los hombres en relaciones homosexuales tienen más probabilidades de tomar riesgos que sus pares heterosexuales.

## Familia

En general, se ha encontrado que los niños y familias con niveles educacionales y económicos más bajos tienen mayores probabilidades de tomar riesgos sexuales. Las familias también pueden

aumentar el riesgo restándole valor a la educación de los niños, fomentando los matrimonios a temprana edad y la maternidad precoz, o desalentando a los jóvenes de obtener información y servicios.

Sin embargo, las familias también pueden proteger a los jóvenes de tomar riesgos. Vivir con los dos padres, tener dinámicas familiares positivas, sentirse apoyado por los padres y otros miembros adultos de la familia y experimentar una supervisión apropiada por miembros adultos de la familia parecen proteger a los jóvenes en la toma de riesgos. Los valores de los padres también influyen sobre los jóvenes; se ha encontrado que los padres y los mayores que se comunican con los jóvenes acerca de sus valores con respecto al sexo los han protegido de una variedad de riesgos. Sin embargo, los resultados de las investigaciones demuestran que está menos claro el impacto de la comunicación sobre sexo y salud reproductiva entre padres y jóvenes en la toma de decisiones de los adolescentes.

### **Instituciones**

La relación con instituciones que apoyan y proporcionan oportunidades a los jóvenes parecen protegerlos de tomar decisiones arriesgadas. Por ejemplo, es menos probable que los jóvenes que se sienten conectados a una organización religiosa tomen riesgos. La conexión a la escuela también es un factor protector, como lo es también el desempeño escolar exitoso y un ambiente sustentador en la escuela. En contraste, algunas instituciones en la comunidad pueden promover la toma de riesgos de los adolescentes. La presencia de industrias del sexo y un acceso amplio a lugares como bares y discotecas pueden aumentar la toma de riesgos de parte de los jóvenes.

Existe evidencia que demuestra que la conexión a organizaciones juveniles también protege a los jóvenes de comportamientos arriesgados. Se sabe que, para adolescentes víctimas de abuso sexual, el acceso a organizaciones que proporcionan actividades de entretenimiento, orientación y servicios protegen a los jóvenes de la toma de riesgos sexuales. Por otra parte, conexiones con otros adultos en la comunidad a través de instituciones sociales, tales como grupos en el vecindario, también son generalmente protectores.

### **Comunidades**

La desorganización o inestabilidad en una comunidad a menudo influye en la toma de riesgos por parte de jóvenes. Altos niveles de desempleo y bajos niveles educacionales, la pobreza, el crimen, la inestabilidad política, la guerra y los altos niveles de migración parecen aumentar la toma de riesgos. La falta de programas, servicios de salud y anticonceptivos, y oportunidades educacionales y económicas en una comunidad también afectan negativamente la toma de decisiones sanas sobre la salud reproductiva por parte de los jóvenes.

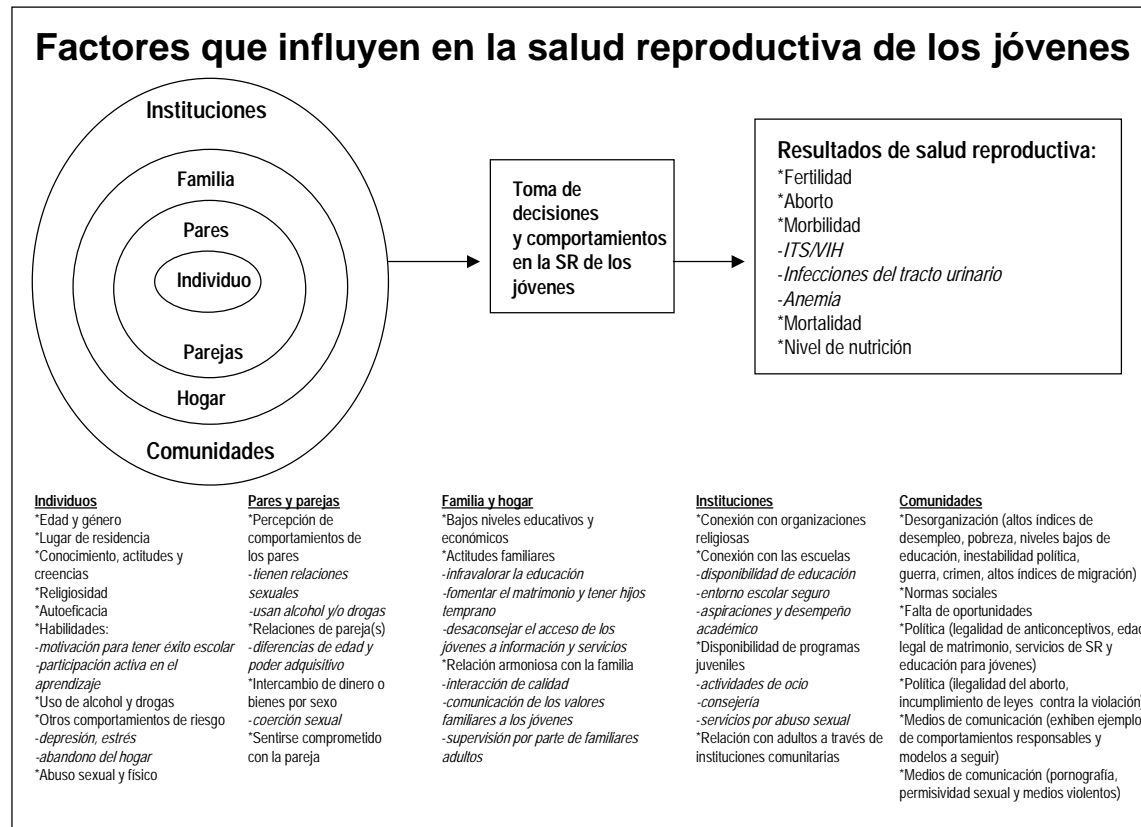
Algunas normas sociales, a pesar de no estar bien documentadas en investigaciones, también parecen influir en que los jóvenes no tomen decisiones sanas sobre su salud reproductiva. La discriminación por género, normas comunitarias que no valoran la educación de los adolescentes, restricciones sobre la movilidad de las niñas y las expectativas culturales de casarse y tener hijos demasiado temprano en la adolescencia pueden impactar negativamente los resultados de salud reproductiva de los adolescentes.

Las políticas de apoyo también pueden proteger a los jóvenes de los riesgos sexuales y de salud reproductiva. Por ejemplo, legalizar la venta de anticonceptivos a los jóvenes e imponer una edad legal mínima para contraer matrimonio pueden ser acciones protectoras. Las políticas que apoyan a la educación y a los servicios de salud para adolescentes también son protectores. A la inversa, la ilegalidad de los abortos y la falta de vigencia/cumplimiento de las leyes contra violaciones y abuso sexual pueden promover resultados negativos de salud reproductiva entre los jóvenes.

Finalmente, los medios de comunicación influyen sobre las normas y valores de la comunidad. Las propagandas y los medios que proporcionan roles positivos y apoyan comportamientos responsables



pueden ser factores protectores. A la inversa, el acceso a la pornografía y a medios sexualmente permisivos o violentos puede aumentar la toma de riesgos entre los jóvenes.



## ...Tres estrategias que promueven la salud reproductiva de los jóvenes

Globalmente, los programas que intentan postergar conductas sexuales arriesgadas y prevenir las ITS han demostrado resultados limitados. Dos factores son en gran parte responsables de esto. En primer lugar, muchas de las evaluaciones han sido de corto plazo y, por lo tanto, no han podido mostrar cambios en comportamientos sexuales y otros resultados de salud reproductiva, tales como las tasas de embarazo y de ITS. En segundo lugar, parece que los programas que se evalúan más a menudo son aquellos que proporcionan información sobre salud sexual y reproductiva y servicios clínicos de salud reproductiva para jóvenes, pero no son suficientes por sí solos para reducir los comportamientos de riesgo de los jóvenes.

Sin embargo, algunos estudios de programas de salud reproductiva para adolescentes sugieren directrices para la planificación de futuros programas de SSRA/SA. En primer lugar, la identificación de antecedentes de riesgo y de factores protectores ha ayudado a los planificadores de programas a identificar a poblaciones objetivo que presentan mayor riesgo de coerción sexual y abuso, de sexo no deseado, de embarazo no intencionado, de ITS y de maternidad insegura. Las evaluaciones también han encontrado que los programas que abordan un espectro más amplio de influencia (en antecedentes protectores) tienden a ser más efectivos en la reducción de los comportamientos de riesgo o el mantenimiento de comportamientos saludables, y tienen mayor probabilidad de tener un impacto a largo plazo.<sup>11</sup> Por ejemplo, algunas evaluaciones han mostrado que los programas de desarrollo juvenil que

<sup>11</sup> Kirby, 1999.

fortalecen las relaciones con la escuela y la familia retardan la iniciación sexual, y reducen los embarazos no deseados y las ITS. De forma similar, los programas que desarrollan habilidades específicas relacionadas con la negociación entre parejas y el uso de preservativos también han conducido a resultados de salud más deseados. Para predecir y atacar los comportamientos arriesgados, investigadores se están concentrando en otros elementos del desarrollo juvenil, tales como utilizar el tiempo de los jóvenes de forma productiva, establecer la presencia de una comunidad compasiva y fomentar un compromiso con el aprendizaje.<sup>12</sup>

Para poder influir efectivamente sobre antecedentes beneficiosos, los programas también deberían iniciar actividades basadas en la promoción de la salud, el cambio social y las teorías de cambio del comportamiento. Las tres estrategias amplias descritas más abajo, cuando se utilizan en forma simultánea, pueden aumentar el impacto sobre la salud de los jóvenes:

- *Aumentar el conocimiento, fomentar actitudes saludables, desarrollar habilidades y formar o cambiar los comportamientos de riesgo de los jóvenes*
- *Mejorar el ambiente social para que los jóvenes sean apoyados en la toma de decisiones saludables, y para que los programas y servicios puedan funcionar*
- *Aumentar el acceso y la utilización de los programas juveniles y los servicios de salud*

### **Estrategia 1: Aumentar el conocimiento, fomentar actitudes saludables, desarrollar habilidades y formar o cambiar los comportamientos de riesgo de los jóvenes**

Esta estrategia, a nivel individual, está dirigida a influir sobre los antecedentes negativos en la toma de decisiones arriesgadas por parte de los adolescentes. Al focalizarse en el fortalecimiento de las características individuales de los jóvenes, se les puede ayudar a tomar decisiones sanas acerca de la salud reproductiva. Este enfoque también puede influir sobre los antecedentes a otros niveles, como por ejemplo cambiando las normas comunitarias, fortaleciendo a las instituciones que apoyan a los jóvenes y fomentando la comunicación efectiva de los adultos con los jóvenes.

- **La transición a la edad adulta requiere conocimientos y habilidades específicas**

Para hacer la transición a la edad adulta, los jóvenes necesitan tener los conocimientos y habilidades que los ayuden a:

- *participar como ciudadanos (como miembros de un hogar, un vecindario y la comunidad, y como trabajadores);<sup>13</sup>*
- *adquirir experiencia en la toma de decisiones;<sup>14</sup>*
- *tomar decisiones basadas en la razón;*
- *evaluar los riesgos y las consecuencias;*
- *evaluar costos y beneficios de las decisiones y acciones; e*
- *interactuar y comunicarse con los pares, parejas y adultos.*

---

<sup>12</sup> Leffert y cols., 1998.

<sup>13</sup> Blum, 1999

<sup>14</sup> Blum, 1999

Este conocimiento y estas habilidades deberían desarrollarse desde una edad temprana, comenzando incluso en la edad preescolar. También deberían ser mejoradas y fortalecidas durante la adolescencia para poder hacer una transición saludable a la edad adulta.

Un volumen creciente de investigaciones indica que los programas de desarrollo juvenil que promueven el conocimiento, las habilidades y otras cualidades personales necesarios para hacer una transición saludable a la edad adulta—en conjunto con información de salud reproductiva y oportunidades para discutir la sexualidad—pueden dar como resultado una amplia gama de consecuencias positivas para la salud.<sup>15 16</sup>

## Sexualidad

La sexualidad incluye no sólo los deseos físicos y sexuales sino también temas de identidad, roles sociales y de género, y relaciones humanas, incluyendo con la familia, los pares y la pareja.

- **Los jóvenes y los adultos encargados de su cuidado necesitan información clara y veraz acerca del sexo**

A medida que los jóvenes van pasando por los cambios físicos relacionados con la reproducción humana, necesitan información acerca de los cambios biológicos que están ocurriendo y oportunidades para hablar sobre la sexualidad de forma segura y abierta.

Con la influencia de los medios de comunicación globales y el cambio de los valores y normas sociales, los jóvenes reciben mensajes inadecuados, mezclados y erróneos acerca del sexo. En muchas sociedades, hablar abiertamente de sexo está prohibido, por lo que muchos jóvenes recurren a amigos, películas, videos y material pornográfico para obtener información. Los padres y encargados adultos a menudo encuentran poco cómodo tratar temas sexuales con sus hijos. Tomando el contexto cultural en consideración, se debería ofrecer información clara y veraz a los jóvenes y a los adultos encargados de su cuidado a través de una variedad de medios y canales.

- **Se deben reforzar las actitudes, intenciones y motivaciones de los jóvenes para prevenir los embarazos y las ITS**

Algunos jóvenes tienen actitudes, intenciones y motivaciones que les llevan a correr riesgos sexuales. Por ejemplo, algunas niñas pueden desear un embarazo porque creen que tener un bebé le dará significado a sus vidas o motivará a sus parejas a casarse con ellas. Otros jóvenes pueden desear actividad sexual sin utilizar preservativos porque creen que los preservativos reducen el placer sexual. Las expectativas culturales pueden motivar a los jóvenes a casarse pronto y tener hijos enseguida. Los programas podrían afectar estas actitudes, intenciones y motivaciones proporcionando orientación o creando grupos pequeños de debate para jóvenes. Estas sesiones pueden ayudar a los jóvenes a examinar críticamente sus actitudes y a cambiar sus intenciones. Por ejemplo, el programa puede ayudar a los jóvenes a examinar los roles tradicionales de género, y ayudarlos a tomar mejores decisiones acerca de qué tipo de relación quieren tener, con quién y cuándo casarse, cuánta educación quieren tener y cuándo quieren tener hijos. Estos programas también pueden ayudar a aumentar la motivación de los jóvenes para evitar embarazos y ETS una vez que examinen las consecuencias.

- **Las actividades de educación en salud pueden afectar todos los factores que influyen en las decisiones de los jóvenes**

Un programa exitoso de SSRA/SA incluye actividades que influyen sobre cómo los jóvenes toman decisiones y actividades que influyen sobre el contexto de la adolescencia. Algunas actividades comunes a los programas de educación de salud reproductiva son:

- *sexualidad, salud reproductiva y educación sobre la vida familiar*

<sup>15</sup> Kirby, 1999.

<sup>16</sup> Leffert y cols., 1998.

- *capacitación en habilidades, incluyendo habilidades para la vida y vocacionales, o habilidades específicas a los comportamientos sexuales, tales como negociación y uso de preservativos*
- *orientación*
- *educación de pares y actividades de extensión*
- *comunicaciones y alcance de los medios de comunicación*
- *referir a los jóvenes a servicios de salud y anticonceptivos*

Estas actividades tienen efectos que interactúan y se sobrepone. Por ejemplo, las comunicaciones y la llegada de los medios de comunicación pueden darle forma a las normas de la comunidad acerca de los jóvenes, y el desarrollo de habilidades puede estimular las oportunidades económicas para los jóvenes.

Las investigaciones han identificado los siguientes elementos programáticos clave de actividades de educación en salud como aquellos que llevan a resultados positivos de salud reproductiva para los jóvenes:<sup>17</sup>

### **Elementos comunes de programas efectivos de educación en salud sexual<sup>18</sup>**

- Presentan un enfoque claro para reducir uno o más comportamientos sexuales que pueden producir embarazos no planeados, ITS o infección por VIH.
- Refuerzan continuamente mensajes claros acerca de comportamientos arriesgados
- Se basan en enfoques teóricos que han demostrado su efectividad para influir otros riesgos relacionados con la salud
- Proporcionan información básica y veraz acerca de los riesgos del coito sin protección y de los métodos para evitar el coito sin protección
- Incluyen actividades que abordan el tema de la presión social sobre los comportamientos sexuales
- Proporcionan modelos y permiten practicar habilidades de comunicación, negociación y negación
- Emplean una variedad de métodos de enseñanza diseñados para involucrar a los participantes y hacerlos personalizar la información
- Incorporan metas de comportamiento, métodos de enseñanza y materiales que son apropiados a la edad, experiencia sexual y cultura de los jóvenes
- Duran el tiempo necesario para completar las actividades importantes
- Usan y capacitan a profesores y educadores de pares que creen en el programa que están implementando

<sup>17</sup> Kirby, 1997; Choi y Coates, 1994; McKaig et al., 1996; y Houvras y Kendall, 1997.

<sup>18</sup> Kirby, 1999b. Esta tabla está basada en el análisis de las evaluaciones efectuadas sobre programas de educación en sexualidad en EE.UU., y pueden tener mayor o menor relevancia en algunos países en vías de desarrollo.

## **Estrategia 2: Mejorar el ambiente social para que los adolescentes puedan tomar decisiones saludables, y los programas y servicios funcionen debidamente**

Mejorar el ambiente social para la salud reproductiva de los jóvenes significa influir en los antecedentes que ocurren entre pares, parejas, familias, instituciones y miembros de la comunidad. Esta estrategia busca cambiar las normas sociales y culturales para apoyar la toma de decisiones sanas por parte de los jóvenes, mejorar los programas y políticas que llegan a los jóvenes, y apoyar a los adultos e instituciones que interactúan y apoyan a los jóvenes.

- **Un ambiente social positivo apoya estilos de vida saludables**

Las relaciones con amigos, parejas y miembros de la familia como también la influencia de la comunidad, la escuela y otras instituciones son importantes a la hora de dar forma a los múltiples resultados de salud. Algunos programas buscan mejorar el ambiente social de la SSRA/SA, incluyendo el fomento de debates críticos sobre normas sociales y culturales que pueden tener un efecto adverso sobre la SSRA/SA, tales como las normas relacionadas a los roles de género. Otros programas pueden tratar de fortalecer instituciones tales como clubes de jóvenes y organizaciones religiosas que llegan y apoyan a los jóvenes, o tratar de desarrollar políticas y programas que proporcionen los servicios que necesitan.

Las comunidades de apoyo pueden tener una influencia positiva. Por ejemplo, la organización comunitaria construye comunidades e instituciones de forma tal que los miembros puedan identificar y resolver problemas y responder a las necesidades. Fomenta el sentido de propiedad y participación, e involucra a los miembros de la comunidad—adultos y jóvenes—en acciones sociales que consideran y abordan las necesidades de salud reproductiva de los jóvenes.

- **El apoyo familiar juega un papel crítico en las decisiones de los jóvenes**

Los padres y otros adultos miembros de la familia juegan un papel crítico en la formación de las aspiraciones y valores de los jóvenes. Incluso cuando los adultos a cargo tienen dificultades para hablar sobre sexo y salud reproductiva con los jóvenes, el apoyo de los adultos puede influir positivamente en los resultados de salud reproductiva de un joven. Se debe estimular a los adultos encargados a valorar la educación de los jóvenes, proporcionar supervisión y apoyo, y comunicarse efectivamente con ellos. Por ejemplo, los programas dirigidos a los padres pueden tener como objetivo ayudarlos a crear una relación armónica con sus hijos, practicando lo que deben decir para demostrar su apoyo a los jóvenes en forma efectiva. Estos programas también pueden exponer a los adultos encargados cómo algunas tradiciones culturales, tales como el matrimonio precoz, tienen efectos negativos en la vida de los jóvenes.

- **Los programas deben identificar y abordar la dinámica de los sistemas sociales de los jóvenes**

Comprender y abordar la dinámica de las relaciones entre los jóvenes y las personas a su alrededor también es un aspecto crítico para mejorar el ambiente de los jóvenes. Muchos programas tienen como objetivo mejorar nuestra comprensión de los sistemas sociales, fortalecerlos y hacer que respondan mejor para apoyar a los jóvenes. Por ejemplo, un programa puede encontrar que algunos jóvenes están en desventaja—tanto en relación con los adultos como con otros jóvenes—debido a las diferencias de edad y experiencia, género, ingreso y educación. Un programa de salud reproductiva para adolescentes puede tratar de mejorar los conocimientos y habilidades de esos jóvenes, pero también puede tratar de influir sobre el comportamiento de aquellos que tienen poder sobre estos jóvenes.

## **Estrategias para crear un ambiente que apoye a los jóvenes**

- Movilizar la acción de la comunidad, particularmente entre los jóvenes.
- Generar respuestas de colaboración sobre SSRA/SA entre los jóvenes, los miembros de la comunidad y las instituciones y organizaciones que trabajan en la comunidad.
- Elevar la conciencia acerca de las necesidades de los jóvenes y los asuntos sociales, culturales, económicos y políticos que contribuyen a sus preocupaciones de salud reproductiva.
- Efectuar campañas de medios de comunicación y mercadeo social.
- Adquirir el apoyo de los interesados y de otros adultos para llevar a cabo debates, actividades y servicios para los jóvenes.
- Abordar los antecedentes que contribuyen a los riesgos de salud reproductiva de los jóvenes, tales como abandono de la escuela, desigualdad de género, matrimonio precoz, corte de los genitales femeninos, la industria del sexo y el consumo de drogas y alcohol.
- Mejorar otros sectores en áreas relacionadas, tales como educación y capacitación vocacional para las mujeres.
- Superar la resistencia en contra de la información y servicios de salud reproductiva para los jóvenes, y asegurarse de que sus precios los hagan accesibles.
- Instituir políticas para promover el acceso a la información, educación y servicios de salud reproductiva. Remover las restricciones que limiten este acceso.
- Apoyar redes y coaliciones que fomentan la abogacía, referencias a otros servicios y cambios sociales más amplios.

### **Estrategia 3: Aumentar el acceso y utilización de los programas y servicios de salud para los jóvenes**

Esta estrategia se concentra en proporcionar oportunidades, programas y servicios que le permitan a los jóvenes ganar acceso a los programas y servicios de salud. Al fortalecer las instituciones que apoyan a los jóvenes, tales como clubes juveniles, instalaciones recreacionales, organizaciones religiosas, escuelas e instalaciones de salud, esta estrategia tiene como objetivo influir sobre la participación de los individuos. La existencia de programas juveniles también puede influir sobre las familias, instituciones y comunidades a medida que estos aumentan la visibilidad de los jóvenes involucrados en actividades positivas, y cambian las actitudes de los adultos hacia ellos.

- **Los programas juveniles pueden afectar la vida de los jóvenes en múltiples niveles**

Muchos programas juveniles tienen como objetivo aumentar el número de jóvenes que participa en actividades que aumentan sus habilidades, forman relaciones positivas con pares y adultos, y proporcionan una salida creativa para su energía. Por ejemplo, los programas juveniles pueden tratar de aumentar las habilidades de los jóvenes, fomentar el activismo en la comunidad, proporcionar deportes, arte u otras actividades creativas, o promover la tutoría de jóvenes por parte de adultos. A nivel individual, estos programas ayudan a desarrollar la autoestima y las habilidades, y estimulan a los jóvenes a tener aspiraciones para su futuro. A nivel interpersonal, fomentan la creación de normas saludables entre los grupos de pares y la interacción positiva entre jóvenes y adultos. A nivel de la comunidad y las instituciones, los jóvenes pueden tener una influencia directa en cambiar el ambiente, si se les estimula a participar abogando por programas y políticas relacionadas a ellos mismos. La

presencia de organizaciones juveniles también puede influir sobre cómo los adultos de la comunidad ven a los jóvenes, y puede ayudar a la comunidad a ver a los jóvenes como un capital importante.

- **La conexión a las escuelas mejora los conocimientos y habilidades de los jóvenes**

Otra estrategia para mejorar la salud reproductiva de los adolescentes consiste en mejorar la calidad y cantidad de educación que ellos reciben. En lugares donde los jóvenes tienen pocas oportunidades educacionales, aumentar el compromiso local y nacional con la educación puede ser una parte importante en el tema de la salud reproductiva de los adolescentes. Además de mejorar la cantidad de educación que reciben los jóvenes, las escuelas pueden mejorar el entorno físico y emocional de la misma escuela. Los programas pueden abordar el tema del acoso sexual en las escuelas, cambiar las políticas escolares que no admiten a estudiantes embarazadas, mejorar la seguridad en las escuelas o fortalecer las actividades extracurriculares.

- **La religiosidad y conexión con organizaciones religiosas pueden influir positivamente sobre los jóvenes**

La adolescencia es una época de cambios rápidos, y las creencias religiosas pueden ayudar a los jóvenes a comprender y procesar los desafíos a los cuales se ven enfrentados. La religiosidad puede tener más que ver con la fuerza de las creencias religiosas de los jóvenes que con su real asistencia o participación en actividades religiosas. Además, la evidencia demuestra que el sentirse conectado a una organización religiosa también puede ayudar a los jóvenes a tomar decisiones saludables. Los programas pueden promover estos vínculos al mismo tiempo que respetan las decisiones individuales en este ámbito.

- **Los servicios de salud permiten a los jóvenes llevar a la práctica sus decisiones saludables**

Proporcionarle servicios de salud a los jóvenes, tales como orientación, anticonceptivos, cuidados maternos y servicios nutricionales, es crucial. Sin estos servicios, los jóvenes pueden tener problemas para ejecutar las decisiones positivas que toman. Para llegar a más jóvenes con servicios de salud necesitamos entender cómo los jóvenes previenen los problemas de la salud reproductiva y buscan tratamiento tanto dentro como fuera de los sistemas formales de entrega de servicios. Muchos estarían de acuerdo en que, para tomar decisiones saludables acerca de enfermedades, es importante ver a un proveedor capacitado de servicios médicos. Sin embargo, los programas de salud reproductiva en general han abordado a mujeres mayores, casadas y a veces a hombres; los jóvenes perciben—a menudo correctamente—que la planificación familiar y las clínicas de ITS no los acogen. Las barreras al acceso a los servicios de salud que enfrentan los jóvenes son numerosas:

- *largas distancias a los lugares de servicio, falta o inseguridad de transporte;*
- *horas de atención inconvenientes;*
- *falta de anonimato;*
- *preocupaciones acerca de la privacidad y confidencialidad;*
- *actitudes y acciones del personal, incluyendo regañones y discursos sobre moralidad;*
- *temor y vergüenza;*
- *costo de los servicios; y*
- *leyes y políticas que hacen difícil servir a los jóvenes.*

Muchos jóvenes dependen de recursos ajenos a los sistemas formales de provisión de servicios de salud. Estos recursos pueden incluir remedios caseros, métodos tradicionales de anticoncepción y abortivos, provisión de anticonceptivos a través de familiares o amigos, abortos clandestinos, y anticonceptivos y medicinas compradas sin receta médica en farmacias o a practicantes tradicionales de la salud. Muchos programas están tratando de aumentar la utilización de los servicios de salud reproductiva por parte de los jóvenes a través de actividades y estrategias que:

- *aumentan los conocimientos de los jóvenes acerca de la disponibilidad de servicios de salud reproductiva;*
- *generan demanda por servicios, por ejemplo, promoviendo servicios a través de pares que trabajan en labores de extensión;*
- *examinan dónde y cómo buscan los jóvenes información y tratamiento, y mejoran los servicios en términos de hacerlos más “fáciles de usar” por los jóvenes.*

Los siguientes puntos son vistos como características de servicios de salud efectivos y “fáciles de usar” por los jóvenes, ya sean proporcionados en una clínica, farmacia, organización de servicios juveniles o por otras vías.<sup>19</sup>

## **Características de los servicios de salud amigables hacia los jóvenes**

### **Características de los proveedores de salud**

- personal especialmente capacitado para trabajar con jóvenes
- respeto hacia los jóvenes
- honrar la privacidad y confidencialidad
- tiempo adecuado para la interacción entre cliente y proveedor
- disponibilidad de educadores de pares

### **Características de las instalaciones de salud**

- espacios separados y horarios especiales
- horarios convenientes
- lugares de atención convenientes
- espacios adecuados y suficiente privacidad
- ambientes confortables/cómodos

### **Características del diseño del programa**

- jóvenes involucrados en el diseño y retroalimentación continua
- clientes espontáneos bienvenidos y citas arregladas con rapidez
- no hay aglomeración de personas y los tiempos de espera son cortos
- precios razonables
- publicidad y reclutamiento que informe y de confianza a los jóvenes
- niños y varones jóvenes bienvenidos y servidos
- amplia gama de servicios disponibles
- disponibilidad de referencias a otros servicios

---

<sup>19</sup> Senderowitz, 1999.



## Otras características positivas

- material educacional disponible en el lugar y para llevar
- disponibilidad de grupos de debate
- es posible recibir anticonceptivos en una visita, aunque se posponen los exámenes pélvicos y de sangre
- formas alternativas de acceso a la información, orientación y servicios fuera de una instalación formal de salud

## ...*Identificación de actividades apropiadas para el programa*

Ahora se tiene una comprensión más clara de los múltiples niveles de influencia sobre la adolescencia y las estrategias amplias que promueven un comportamiento reproductivo saludable entre los jóvenes. Usando esta comprensión como base, se pueden diseñar programas que tengan mayor probabilidad de ser efectivos y, por tanto, que el esfuerzo de una buena evaluación valga la pena. Estos programas:

- *deben definir claramente los resultados de salud que se persiguen*
- *deben identificar los antecedentes protectores y los que aumentan el riesgo que afectan esos resultados*
- *deben usar estrategias programáticas que respondan a más de uno de los antecedentes que impactan los resultados de la salud reproductiva de los adolescentes*

## El modelo lógico

Una forma que ayuda a diseñar la estrategia mientras se tienen estos elementos en mente consiste en usar un **Modelo Lógico**. Los pasos resumidos abajo son:

- Definir las metas de su programa y los resultados esperados (el proceso de definición de metas y resultados será discutido en detalle en el Capítulo 3).
- Identificar los antecedentes que, de acuerdo a la investigación, influyen tanto positiva como negativamente sobre los resultados del comportamiento que el programa busca influir. En muchos lugares puede no existir suficiente investigación para sugerir toda la gama de factores que influyen sobre el comportamiento y la toma de decisiones de los jóvenes. En este caso, se pueden basar las hipótesis en las influencias, ya sea revisando investigaciones que sugieren antecedentes en otros países o usando la experiencia con jóvenes para “intuir” cuáles son los antecedentes que influyen en los resultados de salud. También se debería preguntar directamente a los jóvenes acerca de lo que ellos creen que influye en su toma de decisiones.

### El modelo lógico

El concepto de un “Modelo Lógico” y su importancia en el diseño y evaluación de programas para adolescentes fue introducido por Kirby durante una presentación a la Reunión sobre Salud y Desarrollo de Adolescentes de la OMS en Washington, D.C., que tuvo lugar del 4 al 6 de febrero de 1999. Es una versión simplificada de la estructura lógica que enfatiza que los resultados deberían buscarse basándose en los antecedentes identificados en la investigación.

- Identificar una o más actividades del programa que, basándose en la experiencia propia o en la literatura internacional acerca de lo que funciona, se crea que influyen específicamente en cada antecedente. Por ejemplo:

| <b>Definir las metas de su programa</b>  | <b>Definir el comportamiento deseado de su programa</b>  | <b>Identificar los antecedentes del comportamiento que su programa busca</b>  | <b>Identificar las actividades del programa que Ud. piensa influirán en cada antecedente</b>   |
|--|--|---|--|
| Disminuir las tasas de embarazo y ITS entre los jóvenes de 14 a 19 años en su distrito | <ul style="list-style-type: none"> <li>• Disminuir el sexo prematrimonial</li> </ul>                                       | <ul style="list-style-type: none"> <li>• Normas de la comunidad acerca del sexo prematrimonial y edad apropiada para la iniciación sexual</li> </ul>  | <ul style="list-style-type: none"> <li>• Desarrollar programas de educación para estimular a los adultos a debatir las normas acerca del sexo prematrimonial con los jóvenes</li> </ul>  |
|  | <ul style="list-style-type: none"> <li>• Aumentar el uso de preservativos entre los jóvenes sexualmente activos</li> </ul> | <ul style="list-style-type: none"> <li>• Oportunidades educacionales</li> </ul>   | <ul style="list-style-type: none"> <li>• Iniciar campañas de movilización de la comunidad para cambiar las normas que no valorizan la educación de las niñas</li> <li>• Abogar por la expansión de oportunidades para la educación secundaria</li> </ul>   |
|  | <ul style="list-style-type: none"> <li>• Aumentar la edad de la iniciación sexual</li> </ul>                               | <ul style="list-style-type: none"> <li>• Habilidades individuales para decir “no” al sexo</li> </ul>  | <ul style="list-style-type: none"> <li>• Proporcionar educación de habilidades para la vida en programas escolares de educación en salud, enfatizando cómo decir “no” al sexo</li> </ul>   |
|  |  | <ul style="list-style-type: none"> <li>• Habilidades individuales para usar anticonceptivos</li> </ul>  | <ul style="list-style-type: none"> <li>• Establecer programas de educación de pares para llegar a los jóvenes sexualmente activos</li> </ul>   |
|  |  | <ul style="list-style-type: none"> <li>• Acceso de los jóvenes a preservativos, anticonceptivos y servicios clínicos en forma confidencial</li> </ul> | <ul style="list-style-type: none"> <li>• Estimular el desarrollo de políticas nacionales de salud que apoyen la provisión de servicios a los jóvenes</li> <li>• Incluir representantes juveniles en los comités de asesores clínicos</li> </ul>  |
|  | <ul style="list-style-type: none"> <li>• Aumentar la edad de matrimonio</li> </ul>   | <ul style="list-style-type: none"> <li>• Normas de la comunidad sobre la edad apropiada para el matrimonio</li> </ul>                                 | <ul style="list-style-type: none"> <li>• Establecer servicios que sean confidenciales y acogedores para los adolescentes (educar a los trabajadores de salud, etc.), tanto a nivel nacional como local</li> <li>• Influir en las normas de la comunidad para apoyar un aumento en la edad de matrimonio</li> </ul> |

Utilizando un modelo lógico para guiar el diseño del programa, los supuestos sobre qué influye en los resultados deseados son claros y específicos al contexto. Este modelo aumenta la probabilidad de pensar en forma amplia sobre los factores que influyen en los resultados esperados, y de incluir influencias contextuales, tales como normas comunitarias, las parejas y los pares y las instituciones. Finalmente, las actividades del programa serán diseñadas de tal forma que se vincularán directamente a las influencias que se cree que afectarán el comportamiento deseado.

### ***...Aprendiendo de la experiencia internacional con la programación de la salud reproductiva de los jóvenes***

El diseño de programas de SSRA/SA puede informarse de la experiencia y la evaluación de programas de todo el mundo. Los siguientes “elementos clave” de programas de SSRA/SA fueron recopilados revisando la literatura correspondiente. A pesar de que estos elementos aún no han sido probados sistemáticamente en el terreno, los ofrecemos aquí porque pueden ayudar al diseño e implementación de los programas.

## Elementos clave de programas exitosos de SSRA/SA<sup>20</sup>

- **Planificación estratégica:** Los programas efectivos enuncian claramente los objetivos procesales y de comportamiento antes de que comience el programa, como prerrequisito para medir el éxito.
- **Identificación del grupo objetivo:** Los jóvenes tienen necesidades diversas dependiendo de sus variadas características, tales como edad, situación escolar, estado civil, género, características familiares y experiencia. Al diseñar los programas, es importante identificar el grupo objetivo específico y abordar sus necesidades en conformidad.
- **Evaluación de las necesidades:** Comprender los temas y necesidades específicas de los jóvenes que se espera que participen o reciban servicios del programa asegura que el diseño y contenido del programa se desarrollará apropiadamente.
- **Participación juvenil:** Los jóvenes son los que mejor pueden identificar sus propias necesidades y se sentirán autores/participes de un programa cuando se les incluya en el diseño y la implementación del mismo.
- **Participación en la comunidad:** Los miembros de la comunidad, tales como los responsables de las políticas, profesionales de la salud y líderes religiosos, deberían estar involucrados en la planificación de los programas para asegurar su apoyo y aceptación.
- **Participación de los adultos:** Involucrar a los padres y otros miembros adultos de la familia puede ayudar a asegurar que los programas no enfrenten resistencia y a educar a los padres acerca de temas de salud reproductiva y necesidades de los adolescentes.
- **Protocolos, pautas y estándares:** Políticas operacionales detalladas y específicas que gobiernen cómo un programa debe servir a los jóvenes pueden ayudar a mantener un nivel consistente de calidad, especialmente cuando los proveedores de servicios no conocen bien a la población juvenil.
- **Selección, capacitación y despliegue de personal:** El personal que proporciona servicios a los jóvenes requiere de cualidades, capacitación y supervisión para asegurar que los clientes sean tratados adecuadamente y asegurar el servicio a la clientela.
- **Monitoreo y evaluación:** Recolectar datos para ayudar a los administradores a monitorear el desempeño, a evaluar los resultados y el impacto, y a mejorar las estrategias del programa.

---

<sup>20</sup> Israel y Nagano, 1997; Birdthistle y Vince-Whitman, 1997; Senderowits, 1997a; Senderowitz, 1997b.

## □ Capítulo 3: Desarrollo de un plan de monitoreo y evaluación de SSRA

### Una mirada al capítulo:

- *Define las metas, resultados y objetivos de un programa*
- *Ayuda a establecer la extensión del esfuerzo de monitoreo y evaluación de la SSRA*
- *Ofrece pautas sobre cómo planear y ejecutar una evaluación usando el resto de este manual*

### *...Establecimiento de metas, resultados y objetivos para los programas de salud reproductiva de los jóvenes*

Esta sección trata las metas, resultados y objetivos de un programa de salud reproductiva de adolescentes. Cada uno es una expresión diferente de los resultados de salud reproductiva que el programa está tratando de alcanzar. Ellos forman la base del esfuerzo de M&E.

#### **Las metas definen el impacto global que los programas esperan tener**

Una meta indica el impacto que un programa pretende tener sobre la población objetivo. La población objetivo es el grupo específico de individuos a los cuales el programa quiere afectar, y puede incluir jóvenes y también los adultos proveedores de servicios, profesores, miembros de la familia y de la comunidad que interactúan con los jóvenes. Los programas de SSRA a menudo tienen la meta general de mejorar la salud reproductiva de los jóvenes. Las metas pueden indicarse más específicamente dependiendo de las necesidades de salud reproductiva de la población joven.

#### **Los efectos del programa se relacionan a los resultados específicos que el programa pretende lograr**

Los efectos esperados del programa están relacionados a las metas establecidas, tales como la disminución de las tasas de ITS o mejoras en la situación nutricional. Para producir estos resultados, los programas se concentran en cambios intermedios de comportamiento, tales como retrasar la iniciación sexual, aumentar el uso de preservativos o anticonceptivos, o aumentar la lactancia materna. Los programas pueden establecer resultados de corto, mediano y largo plazo, tal como se detalla a continuación.

#### **Terminología**

Las personas que trabajan en evaluación usan diferentes términos para describir lo que hacen. La terminología actual de evaluación a menudo se interpreta de maneras distintas de acuerdo a los diferentes ambientes y, a veces, los evaluadores pasan demasiado tiempo debatiendo qué término es mejor. En este manual usamos términos y conceptos cuya intención es reflejar las etapas y componentes de los programas juveniles, tal y como se implementan en el campo. Los hemos definido de forma tal que esperamos sean comprensibles y accesibles a aquellos que no tienen experiencia en el área de investigación.

| <b>Meta: Mejorar la salud reproductiva de las mujeres jóvenes</b>                                    |   |   |
|--|---|---|
| <b>Resultado a corto plazo</b>   | <b>Resultado a mediano plazo</b>  | <b>Resultado a largo plazo</b>  |
| Mejorar la calidad de la interacción entre los padres y las jóvenes entre 10-19 años en su distrito. | Aumentar la edad promedio de la iniciación sexual entre las jóvenes de 14-19 años en su distrito. | Disminuir las tasas de embarazo entre las jóvenes de 14-19 años en su distrito. |

### **Los objetivos son afirmaciones explícitas y medibles de los resultados del programa**

Los objetivos son afirmaciones explícitas y medibles de los resultados del programa. Hay dos tipos de objetivos, a nivel de población y a nivel de programa.

Los objetivos a nivel de población indican los resultados que se pretenden alcanzar dentro de la población objetivo, y están directamente relacionados a los resultados identificados en el programa. Describen el impacto que quiere lograr el programa en la población juvenil a la que aspira llegar, influir o servir. Por ejemplo:

- *Aumentar en un año la edad promedio de la iniciación sexual entre los jóvenes de 14 a 19 años en su distrito*
- *Aumentar el porcentaje de jóvenes entre 14 y 19 años en el distrito del programa que están activamente involucrados en organizaciones juveniles que ofrecen actividades de entretenimiento*

Los objetivos a nivel de programa indican los resultados que se esperan en cuanto a la estructura, administración u operaciones de un programa. Describen las actividades que se llevarán a cabo para lograr el impacto que el programa quiere tener. Por ejemplo:

- *Capacitar a 30 educadores pares para proporcionar orientación de calidad a los jóvenes cada seis meses.*

### ***...Medición de los objetivos***

La forma de conceptualizar y expresar los objetivos y sus mediciones enmarcará las acciones a tomar.

### **La medición de un objetivo debería establecerse en términos de las metas**

Las metas fijan el nivel de los objetivos que se desean alcanzar en un tiempo especificado.<sup>21</sup> Las metas pueden ser cuantitativas (numéricas) o cualitativas (descriptivas), dependiendo de la naturaleza de la actividad y el indicador que se elija para medirlas.<sup>22</sup> Las metas pueden expresar cantidad (cuánto), calidad (cuán bien) o eficiencia (menor costo por resultado producido).

<sup>21</sup> Las metas son estimados cuantitativos que se usan con el propósito de presupuestar, planificar y darle seguimiento a los cambios en los resultados. No deberían entenderse como cuotas o usarse como base para forzar a un individuo a aceptar servicios tales como la anticoncepción, que sean inconsistentes con sus creencias morales, filosóficas o religiosas. Las metas no deberían usarse como base de compensación de los proveedores de servicios. Todos los programas de salud reproductiva de los jóvenes deberían salvaguardar los derechos, la salud y el bienestar de todos los individuos que toman parte en el programa.

<sup>22</sup> Los indicadores se discuten en detalle en el Capítulo 5.

### **Las metas de los objetivos a nivel de población deberían ser definidas refiriéndose a la información base**

La información base describe la situación existente en una comunidad antes de que se lleve a cabo una intervención. La información base es importante porque proporciona puntos de comparación para medir y ver si se han logrado los objetivos. Si no se dispone de información base, quizá se necesite recolectar información sobre la población objetivo y sus necesidades antes de que comience el programa, ya que eso proporcionará mediciones iniciales que pueden ser la base para evaluaciones de resultados o impactos que el programa realice más adelante.

La fuente de la información base podría ser:

- *una encuesta de los jóvenes antes de la intervención;*
- *datos documentando experiencias previas de programas juveniles;*
- *medidas externas recolectadas por otra organización, agencia gubernamental o donante tal como datos sobre utilización de servicios gubernamentales de salud;*
- *información sobre salud reproductiva de los jóvenes obtenida de una encuesta nacional, tal como una Encuesta Demográfica y de Salud (EDS);*
- *el juicio profesional de aquellos que trabajan con jóvenes.*

Por ejemplo, la experiencia previa con programas puede decir que sólo el 5% de los jóvenes buscan servicios de orientación de educadores pares en las escuelas donde funciona el programa. Pero se sabe que una organización asociada en un distrito vecino ha determinado que 8% de los jóvenes buscaron orientación. Con esta información base se podría determinar que el objetivo del programa debería ser “aumentar el porcentaje de jóvenes entre 10 y 19 años que buscan servicios de orientación de educadores pares a 10% en un año”.

### **Las metas de los objetivos a nivel de programa deberían ser definidas por la experiencia en los programas**

Para determinar las metas de los objetivos del programa, tales como el número de adolescentes educadores que deberían ser capacitados, hay que referirse a la experiencia y los recursos con que cuentan los programas. Por ejemplo, se puede determinar que para llegar a 500 jóvenes en la población objetivo, se necesita capacitar a 40 educadores pares. Pero si el presupuesto sólo alcanza para una sesión de capacitación cada seis meses, y la experiencia demuestra que capacitar a 20 educadores pares al mismo tiempo es lo más efectivo, podría fijarse como meta capacitar a 20 educadores pares cada seis meses.

## Cómo traducir metas y efectos en objetivos

|   |  |   |
|---|--|---|
| <p><b>Meta:</b></p> <ul style="list-style-type: none"> <li>Mejorar la salud reproductiva de los jóvenes entre 10 y 19 años en su distrito.</li> </ul>   | <p>Para traducir a resultados, describir los resultados específicos que el programa busca lograr.</p>  | <p><b>Resultados:</b></p> <ul style="list-style-type: none"> <li>Aumentar el uso de preservativos entre los jóvenes sexualmente activos entre 15 y 19 años en el distrito del programa.</li> <li>Aumentar el número de jóvenes sexualmente activos menores de 19 años que tratan el uso de preservativos con sus parejas.</li> </ul>  |
| <p><b>Resultados:</b></p> <ul style="list-style-type: none"> <li>Aumentar el uso de preservativos entre los jóvenes sexualmente activos entre 15 y 19 años en el distrito del programa.</li> <li>Aumentar en el distrito del programa, el número de jóvenes sexualmente activos menores de 19 años que tratan el uso de preservativos con sus parejas.</li> </ul> | <p>Para traducir los resultados a objetivos de población, referirse a la información base: La EDS más reciente encontró que el 5% de los jóvenes entre 15 y 19 años en su distrito usaron preservativos durante el primer coito. Una encuesta base efectuada por la organización encontró que el 15% de los jóvenes sexualmente activos entre 15 y 19 años han hablado alguna vez sobre el uso de preservativos con sus actuales parejas sexuales.</p> | <p><b>Objetivos a nivel de población:</b></p> <ul style="list-style-type: none"> <li>Aumentar en el distrito del programa el uso de preservativos al momento del primer coito entre los jóvenes de 10 a 15 años a 10% dentro de un año.</li> <li>Aumentar en el distrito del programa el número de jóvenes sexualmente activos menores de 19 años que alguna vez han tratado el uso de preservativos con sus actuales parejas sexuales a 25% dentro de un año.</li> </ul> |
| <p><b>Resultados:</b></p> <ul style="list-style-type: none"> <li>Aumentar en el distrito del programa el uso de preservativos entre los jóvenes sexualmente activos de 15 a 19 años.</li> <li>Aumentar el número de jóvenes sexualmente activos menores de 19 años que tratan el uso de preservativos con sus actuales parejas sexuales.</li> </ul>               | <p>Para traducir los resultados a objetivos de programa hay que describir las actividades que se llevarán a cabo para lograr los efectos.</p>  | <p><b>Objetivos a nivel de programa:</b></p> <ul style="list-style-type: none"> <li>Capacitar a 25 educadores pares para realizar reuniones que aumenten los conocimientos de los jóvenes entre 15 y 19 años sobre el uso de preservativos y sobre cómo negociar.</li> <li>Efectuar 30 sesiones con jóvenes entre 15 y 19 años, para aumentar sus conocimientos acerca del uso de preservativos y cómo negociar.</li> </ul>   |

### **El monitoreo y evaluación requiere comprender mediciones e indicadores**

La *medición* es el uso de métodos y procedimientos para la observación y evaluación sistemática.<sup>23</sup> Se utilizan una variedad de métodos y procedimientos para recolectar información acerca de un programa y su población objetivo.<sup>24</sup>

<sup>23</sup> Green y Lewis, 1986.

<sup>24</sup> La recolección de datos se trata en detalle en el Capítulo 4.



Para medir el funcionamiento de un programa y los efectos que está teniendo en la población objetivo se usan *indicadores*. Un indicador es una medida de los objetivos y actividades del programa.<sup>25</sup> Los cambios en los indicadores demuestran que el programa está funcionando y qué consecuencias—positivas y negativas—tiene el programa en la población objetivo.

Se recolecta información acerca de algunos objetivos—tanto a nivel de programa como a nivel de población—para medir si las actividades del programa se están implementando, para medir la calidad de la implementación del programa, para saber en qué medida se está utilizando el programa; o qué cambios están ocurriendo en la población objetivo. En general, la información recogida durante una evaluación de proceso medirá los objetivos a nivel de programa. La información recogida durante una evaluación de resultados o de impacto medirá los objetivos a nivel de población. Para medir los cambios en los objetivos, se compara la información base con los datos recogidos después de que el programa ha estado operando por algún tiempo.

El monitoreo, la evaluación de procesos, la evaluación de resultados y la evaluación de impacto involucran los siguientes pasos:

- *estar de acuerdo en la amplitud y objetivos del plan de M&E con los interesados,*
- *seleccionar los indicadores,*
- *recoger información acerca de esos indicadores en forma sistemática y consistente,*
- *analizar la información recogida,*
- *comparar los resultados con las metas y objetivos iniciales del programa, y*
- *compartir los resultados con los interesados, incluyendo a los jóvenes.*

### **Un plan sólido de M&E debería usar indicadores para medir tanto los objetivos a nivel de población como a nivel de programa**

Los objetivos a nivel de población están relacionados más directamente a los resultados de salud sexual y reproductiva que el programa quiere lograr. Sin embargo, a menudo algunos resultados son difíciles de medir porque tratan temas delicados, tales como si los jóvenes practican el sexo o no. Siempre se debería tratar de medir los objetivos a nivel de población en relación a los resultados intermedios de comportamiento, pero puede ser difícil lograrlo.

Medir los objetivos de corto plazo relacionados a los factores de riesgo y protección que el programa piensa que influyen en el comportamiento de los jóvenes es importante por dos razones. En primer lugar, en ausencia de cambios observables del comportamiento a largo plazo, el logro de los objetivos de corto plazo es una buena señal de que el programa está produciendo resultados. En segundo lugar, medir los objetivos de corto plazo también ayuda a probar las hipótesis acerca de qué factores influyen en el comportamiento y en la toma de decisiones de los jóvenes. Esta información puede ayudar a determinar si está funcionando o no la estrategia del programa para influir en el comportamiento que produce los resultados de salud reproductiva a largo plazo.

Medir los objetivos a nivel de programa es importante para entender cómo está funcionando el programa. Los objetivos a nivel de programa se miden como parte de una evaluación de procesos y proporcionan información acerca de cómo está funcionando el programa. Una evaluación de procesos

---

<sup>25</sup> Los indicadores se tratan en detalle en el Capítulo 4.

puede proporcionar conocimientos acerca de porqué el programa tiene impacto o no lo tiene, y es importante si se planea aumentar la escala del programa o replicar las estrategias que usa el programa.

| <b>Hoja de trabajo 3.1</b><br><b><i>Identificación de metas, resultados, contexto y objetivos del programa</i></b>  |  |
|---|--|
| 1. ¿Cuáles son las metas del programa?  |  |
| 2. ¿Cuáles son los resultados de corto, mediano y largo plazo que el programa espera lograr?  |  |
| 3. ¿Cuáles son los objetivos de corto plazo a nivel de población que el programa espera lograr (incluyendo objetivos relacionados a factores antecedentes)? |  |
| 4. ¿Cuáles son los objetivos a mediano plazo a nivel de población que el programa espera lograr?  |  |
| 5. ¿Cuáles son los objetivos a nivel de programa?<br>¿Cómo se lograrán los objetivos indicados más arriba a nivel de población?                             |  |
| 6. ¿Qué actividades serán implementadas por el programa?  |  |
| 7. ¿Quiénes son los interesados en el programa?   |  |
| 8. ¿Cómo podría el contexto local político o cultural afectar al programa?  |  |
| 9. ¿Qué impacto tendrán las condiciones económicas existentes en la implementación del programa o en la participación de los jóvenes?                       |  |

### ***...Definiendo el alcance de un esfuerzo de M&E***

**Alcance** se refiere al nivel de actividad que se realizará en un esfuerzo de supervisión y evaluación. El alcance del esfuerzo de M&E está determinado por varios factores. Hay que hacerse seis preguntas clave:

- *¿Qué debería ser monitoreado y evaluado?*
- *¿Cuándo deberían ser monitoreados y evaluados los programas de SSRA/SA?*

- *¿Cuánto costará el M&E?*
- *¿Quién debería estar involucrado en el M&E?*
- *¿Quién debería llevarse a cabo la evaluación?*
- *¿Dónde debería efectuarse el M&E?*

Cada una se trata a continuación.

### **¿Qué debería ser monitoreado y evaluado?**

- **El M&E puede medir cada etapa del programa: diseño, desarrollo y funcionamiento de los sistemas e implementación**

Después de establecer las metas, objetivos y actividades, el próximo paso es tomar decisiones respecto del M&E en las etapas del desarrollo del programa. Estas etapas son: (1) diseño, (2) desarrollo y funcionamiento de sistemas, e (3) implementación.

El esfuerzo de M&E puede medir cada una de estas etapas para determinar cómo está funcionando el programa y su impacto sobre la población objetivo. Se puede revisar cada etapa para generar ideas y opciones para los esfuerzos de M&E.

- **El diseño del programa se mide con una evaluación de procesos**

Una evaluación de las necesidades programáticas de la comunidad a menudo sirve de base para el diseño de programas. El proceso de diseño de programas involucra desarrollar una estrategia o enfoque sistemático para abordar las necesidades de la comunidad, identificar las acciones y actividades que se requieren para implementar la estrategia, e identificar los recursos adicionales necesarios (ya que la evaluación de necesidades identifica los recursos que ya están presentes), para llevar a cabo las actividades. Evaluar cuán bien se ha diseñado un programa es un aspecto de la evaluación de proceso porque el diseño del programa afecta su éxito. Documentar los problemas en el diseño del programa ayudará a explicar porqué el programa no logró sus objetivos y—si es exitoso—qué elementos claves del diseño contribuyeron al éxito. Entonces, esos elementos podrán utilizarse para expandir o reproducir el programa en el futuro.

- **El desarrollo y funcionamiento de los sistemas se mide a través del monitoreo y evaluación de procesos**

El *desarrollo de sistemas* involucra la creación de un sistema de administración y apoyo para ejecutar el programa. Los sistemas de apoyo incluyen los sistemas de información gerencial (SIG), los sistemas de administración financiera, sistemas de personal y bienes, y sistemas logísticos. Llevar a cabo actividades preparatorias, tales como reclutamiento y capacitación del personal, desarrollo de planes de estudio, de pautas de operación y desarrollo de materiales de información, educación y comunicación (IEC) o comunicación de cambios de comportamiento (CCC), es una parte importante del desarrollo de sistemas.

El *funcionamiento de los sistemas* involucra el desempeño continuo de los sistemas usados para operar el programa e incluye temas tales como: forma en que se toman las decisiones dentro del programa, si los canales de comunicación internos y externos funcionan bien, cuán bien se efectúa la coordinación entre los programas regionales y la sede, si la capacitación y supervisión están asegurando un desempeño de alta calidad, y cuáles son los términos de referencia y de desempeño para el personal.

Si se puede documentar como funcionan los sistemas del programa, esto ayudará a explicar por qué el programa funciona o no. Para determinar cómo están operando los sistemas de un programa, el monitoreo y evaluación de proceso deberían:

- *documentar el desarrollo de sistemas de apoyo y determinar si realmente operan una vez que comienza la implementación del programa;*
  - *evaluar el desempeño de los sistemas de apoyo; y*
  - *medir cuán efectivas son las actividades preparatorias en alistar al personal del programa para su implementación.*
- **La implementación se mide a través del monitoreo, evaluación del proceso y evaluación del resultado/impacto**

La *implementación* es el proceso de llevar a cabo las actividades del programa con la población objetivo y proporcionarles servicios. Es la ejecución de las actividades planeadas. Por ejemplo, las actividades de un centro juvenil pueden incluir contratar y capacitar al personal y a los voluntarios en sesiones educativas realizadas en el centro, involucrar a los jóvenes en la creación de actividades de entretenimiento y proporcionar servicios de orientación a los jóvenes.

El monitoreo y evaluación de proceso revelan cómo se está dando la implementación del programa. La evaluación de resultados y de impacto ayudan a determinar si el programa está logrando sus objetivos midiendo los cambios en los resultados en la población objetivo. Esta información, en conjunto, debería ayudar a explicar por qué el programa está—o no está—logrando sus objetivos y contribuye a la comprensión de los resultados del programa.

- **Las metas, objetivos y actividades del programa establecen el alcance de lo que será monitoreado y evaluado**

Al identificar todas las actividades que el programa ha realizado en las etapas de diseño, sistemas e implementación se definen las posibilidades del esfuerzo de M&E. En cada etapa, el desarrollo de las actividades deben ser monitoreadas y/o evaluadas. El cuadro que sigue ilustra como las actividades que se han llevado a cabo en cada etapa de un programa de educación de pares pueden ser monitoreadas y evaluadas. Identificar las actividades en cada etapa y definir las posibilidades de M&E es el primer paso para determinar el alcance.

| Actividades  | Monitoreo y evaluación de procesos  | Evaluación de resultados e impacto  |
|--|---|---|
| <b>Etapa de diseño</b>   |   |   |
| <ul style="list-style-type: none"> <li>Determinar si los adolescentes educadores son una forma efectiva de llegar a la población objetivo</li> </ul>   | <ul style="list-style-type: none"> <li>¿Se consultó a los jóvenes de la población objetivo acerca de si pensaban que los adolescentes educadores serían efectivos?</li> </ul>                                 | N/A   |
| <b>Etapa de desarrollo de sistemas y funcionamiento</b>  |   |   |
| <ul style="list-style-type: none"> <li>Desarrollar planes de estudio para capacitar a los adolescentes educadores</li> <li>Reclutar, seleccionar y capacitar a los educadores pares</li> </ul> | <ul style="list-style-type: none"> <li>¿Cuántos adolescentes educadores son reclutados, seleccionados y capacitados?</li> <li>¿Qué calidad tiene la capacitación entregada a los educadores pares?</li> </ul> | N/A   |
| <b>Etapa de implementación</b>   |   |   |
| <ul style="list-style-type: none"> <li>Adolescentes educadores proporcionan orientación tres tardes a la semana en cinco clínicas de salud</li> </ul>  | <ul style="list-style-type: none"> <li>¿Cuántos jóvenes son orientados por educadores pares?</li> <li>¿Qué calidad tiene la orientación proporcionada por los educadores pares?</li> </ul>                    | <ul style="list-style-type: none"> <li>¿Ocurren cambios en el conocimiento, actitudes y comportamiento de los jóvenes orientados por educadores pares?</li> </ul> |

- Cómo se planea utilizar la información de M&E que dará forma al monitoreo y evaluación**

El uso que se le dará a la información de M&E ayudará a determinar el alcance del esfuerzo de M&E. Posibles usos incluyen:

- Si el propósito principal es asegurar que las actividades del programa se lleven a cabo tal como están planeadas, se necesita dar seguimiento a los cambios de los objetivos a nivel de programa, mediante un sistema efectivo de monitoreo.
- Si interesa evaluar cuán bien se están llevando a cabo las actividades y hacer mejoras a medida que sea necesario durante el transcurso de la implementación del programa, se debe realizar una evaluación de proceso.
- Si interesa determinar que los indicadores de resultados estén cambiando en la población objetivo del programa, se debe realizar una evaluación de resultados. Si se dispone de más recursos e interesa mostrar cuántos de los cambios observados en los indicadores de resultados se deben al programa, se necesita hacer una evaluación de impacto.
- Puede que sea necesario responder a los requisitos de los donantes en cuanto a que el programa realice evaluaciones de resultados o impactos.
- Cómo se está desempeñando el programa y sobre qué resultados influye ayudará a decidir si se continúa, si se cambia o se expande la estrategia del programa.

**Razones para monitorear y evaluar:  
Diferentes necesidades para diferentes actores**

| <b>Administradores y personal del programa</b>   | <b>Agencias financieras y elaboradores de políticas</b>   | <b>Comunidades y juventud</b>   |
|--|---|---|
| <b><i>Lo que mide el M&amp;E</i></b>   | <b><i>Lo que mide el M&amp;E</i></b>  | <b><i>Lo que mide el M&amp;E</i></b>  |
| <ul style="list-style-type: none"> <li>• calidad de las actividades y/o servicios</li> <li>• por qué algunos lugares son menos exitosos</li> <li>• capacidad y técnicas de M&amp;E</li> <li>• cobertura del programa</li> </ul>  | <ul style="list-style-type: none"> <li>• evidencia del logro de los objetivos</li> <li>• resultados e impacto del programa</li> <li>• costo-eficiencia del programa</li> <li>• datos sobre salud reproductiva de los jóvenes</li> </ul>   | <ul style="list-style-type: none"> <li>• comportamientos de los jóvenes en cuanto a salud reproductiva</li> <li>• necesidades de los jóvenes</li> <li>• cómo se gastan los fondos del programa</li> <li>• el proceso e impacto de la participación de la comunidad</li> </ul>   |
| <b><i>Lo que identifican los resultados de M&amp;E:</i></b>  | <b><i>Lo que identifican los resultados de M&amp;E:</i></b>   | <b><i>Lo que identifican los resultados de M&amp;E:</i></b>   |
| <ul style="list-style-type: none"> <li>• prioridades para la planificación estratégica</li> <li>• necesidades de capacitación y supervisión</li> <li>• cómo mejorar los informes a las agencias financieras</li> <li>• retroalimentación de los clientes</li> <li>• por qué el programa no está logrando lo que se planeó</li> </ul> | <ul style="list-style-type: none"> <li>• prioridades para el financiamiento estratégico del programa</li> <li>• qué programas califican para asistencia de donantes</li> <li>• qué “mejores prácticas” deberían requerir los donantes de los programas juveniles</li> <li>• los impactos de la asistencia de donantes</li> </ul>  | <ul style="list-style-type: none"> <li>• los beneficios reales y potenciales de los programas juveniles</li> <li>• la necesidad de nuevos y mejores servicios para los jóvenes</li> <li>• recursos de la comunidad que se pueden usar para apoyar los programas de SSRA</li> <li>• la necesidad de apoyo local para temas y para actividades de SSRA</li> </ul> |
| <b><i>Qué decisiones guían los resultados:</i></b>   | <b><i>Qué decisiones guían los resultados:</i></b>  | <b><i>Qué decisiones guían los resultados:</i></b>  |
| <ul style="list-style-type: none"> <li>• distribución de recursos</li> <li>• reproducción y aumento de las intervenciones</li> <li>• recaudación de fondos</li> <li>• motivar al personal</li> <li>• abogacía de políticas</li> <li>• movilización de la comunidad</li> </ul>  | <ul style="list-style-type: none"> <li>• cuánto financiamiento se debería asignar a la SSRA</li> <li>• qué tipos de programas juveniles deberían ser financiados</li> <li>• qué enfoques de programas deberían ser presentados como modelos</li> <li>• nuevos objetivos estratégicos, actividades o paquetes de resultados</li> <li>• reproducción y aumento de los programas exitosos</li> </ul> | <ul style="list-style-type: none"> <li>• en qué grado deberían participar los miembros de la comunidad y los jóvenes en los programas de apoyo</li> <li>• cómo coordinar de mejor forma las acciones de la comunidad para abordar la SSRA</li> <li>• cuánto y qué tipo de recursos locales deberían ser asignados a la SSRA</li> </ul>                          |

## **¿Cuándo deberían ser monitoreados y evaluados los programas de SSRA?**

El monitoreo y evaluación de proceso deberían ocurrir a lo largo de la vida del programa. La información recogida se puede usar para asegurar que se están cumpliendo los objetivos, para mejorar el desempeño del programa y proporcionar retroalimentación y apoyo al personal y a los participantes del programa.

Las evaluaciones de resultados y de impacto generalmente se hacen cerca del fin del programa, a pesar de que a menudo usan información sobre condiciones iniciales que se han recogido al comienzo del programa. A menudo, una evaluación de impacto tiene que ser diseñada dentro de un programa desde el comienzo o no se tendrá el tipo de información inicial necesaria para medir los cambios en los resultados, y atribuirlos al programa.

Es muy importante no efectuar una evaluación de resultados o impactos prematuramente. Para algunos resultados esperados, tales como cambios en los comportamientos arriesgados, las actividades del programa necesitan efectuarse por algún tiempo, quizás años, antes que puedan observarse cambios en la población objetivo. En este caso, las evaluaciones de resultados o de impacto pueden efectuarse antes del final, pero después de que el programa haya estado en pleno funcionamiento por algún tiempo.

Para determinar cuándo efectuar evaluaciones, hay que basarse en los objetivos del programa, las necesidades de información acerca del programa de los diversos interesados, el conocimiento que se tenga del programa, los recursos disponibles y el juicio de los administradores.

El punto en el programa en que se comience el esfuerzo de M&E determinará el tipo de monitoreo y evaluación que se puede llevar a cabo.

### **Si se comienza el M&E muy tarde...**

- Puede no contarse con la información base acerca de la situación de la población objetivo antes del comienzo del programa;
  - La información recogida será menos significativa;
  - La información no ayudará para hacer mejoras en la estrategia del programa; y
  - Los resultados de la evaluación serán menos concluyentes acerca de si ocurrieron cambios en los resultados o si los cambios se pueden atribuir a las actividades del programa.
- 
- **Comenzar el M&E al comienzo del programa es lo ideal**

El monitoreo y evaluación debería ser planificado—y empezado—al comenzar un nuevo programa. La planificación temprana permite definir el esfuerzo de M&E basado en los objetivos y actividades, y permite ser estratégico acerca de lo que se planea medir. También permite encontrar la información existente y recolectar información base en el momento ideal—cuando el programa está empezando. Esto permitirá llevar a cabo evaluaciones de resultados o de impacto con mayor facilidad y aumentará la capacidad de medir el verdadero impacto. Comenzar el monitoreo y evaluación de proceso temprano también permite utilizar los resultados de M&E para mejorar el programa mientras está siendo implementado. Finalmente, comenzar temprano permite asegurar que los costos de M&E estén adecuadamente considerados en el presupuesto.

- Algunas actividades se pueden medir aun cuando el M&E comience a mitad del programa

Se puede dar el caso en que se necesite un plan de M&E una vez comenzado el programa. Si se comienza el esfuerzo de M&E en la mitad del programa, su alcance probablemente será limitado. Pero puede ser posible llevar a cabo una evaluación de resultados, para la cual probablemente se tendrá que utilizar información base recogida después del comienzo del programa. A pesar de que los resultados pueden no ser tan claros y sólidos, todavía pueden ser útiles. Aunque un SIG se puede establecer a mitad del programa para dar seguimiento a los resultados del monitoreo y evaluación de proceso, éste será menos útil que si se lanza al comienzo.

- Si el M&E se comienza hacia el final del programa, se pueden medir menos actividades

Puede ser que algunos administradores de programa no piensen en lo que van a monitorear o evaluar hasta que el programa esté casi terminado. Si se comienza el esfuerzo de M&E al final del programa, las posibilidades de llevar a cabo un buen M&E son muy limitadas. En primer lugar, no es muy útil establecer un sistema de monitoreo al final del programa. En segundo lugar, aunque se pueden evaluar las actividades del programa en retrospecto (por ejemplo, solicitando retroalimentación de los participantes e interesados cuando el programa ya está en funcionamiento), los resultados pueden ser sesgados. Finalmente, a pesar de que una evaluación de resultados es posible, ésta tendrá que depender de la utilización de estándares externos—estimaciones sobre la posible situación en la comunidad antes de que se llevara a cabo la intervención—como datos de comparación. Estos estándares pueden o no reflejar verazmente el conocimiento, las actitudes y comportamientos de la población objetivo antes de que empezara el programa, limitando así la posibilidad de demostrar cambios en los resultados.

### Flujo de un esfuerzo de M&E comenzado al comienzo del programa

| Etapa del programa | Monitoreo  | Proceso  | Resultado/Impacto  |
|--------------------|--|--|--|
| <b>Temprana</b>    | <ul style="list-style-type: none"> <li>• Establecer un sistema de monitoreo (SIG), identificar indicadores e instrumentos, planificar el seguimiento del programa, análisis de datos, informes.</li> </ul>   | <ul style="list-style-type: none"> <li>• Evaluar el desarrollo de los sistemas y su funcionamiento, incluyendo la capacitación y supervisión del personal. Proporcionar retroalimentación temprana. Evaluar si el programa responde a los jóvenes o si necesita nuevos elementos.</li> </ul> | <ul style="list-style-type: none"> <li>• Identificar objetivos e indicadores. Medir las condiciones iniciales. Crear un plan de evaluación de resultados o impacto.</li> </ul>   |
| <b>Media</b>       | <ul style="list-style-type: none"> <li>• Evaluar el SIG y los datos. Si el sistema original es inadecuado o si el programa agrega nuevos componentes. Modificarlo. Si el programa no se está desempeñando como estaba planificado, iniciar una evaluación de proceso.</li> </ul> | <ul style="list-style-type: none"> <li>• Efectuar una evaluación de proceso más formal de mediano plazo para evaluar la calidad del desempeño del programa. Determinar la cobertura, o si el programa está llegando a la población objetivo.</li> </ul>                                      | <ul style="list-style-type: none"> <li>• Tomar medidas a medio camino. Analizar las medidas de resultados de corto plazo, tales como cambios en el conocimiento, aumento en el uso de programas y cambios en los factores contextuales. Proporcionar retroalimentación al programa.</li> </ul> |



|               |   |  |   |
|---------------|---|--|---|
| <b>Tardía</b> | <ul style="list-style-type: none"> <li>Analizar datos del sistema de seguimiento para ver si el programa se ejecutó como estaba planeado. Preparar y presentar informes.</li> </ul> | <ul style="list-style-type: none"> <li>Analizar mediciones del final del programa. Determinar qué se hizo para mejorar la calidad de la implementación del programa. Hacer recomendaciones para repetir o expandir el programa.</li> </ul> | <ul style="list-style-type: none"> <li>Hacer seguimiento al final del programa. Examinar la evidencia de los cambios de resultados. Dependiendo del diseño del estudio, efectuar un análisis de impacto para ver si los resultados se pueden atribuir a las actividades del programa. Informar a los donantes y otros interesados.</li> </ul> |
|---------------|---|--|---|

### ¿Cuánto costará el monitoreo y la evaluación?

- Los recursos financieros disponibles influirán en el nivel de evaluación que se emprenda

Los administradores de programas deben determinar si el tiempo, esfuerzo y costo de una evaluación se justifican en vista de los beneficios esperados. La disponibilidad de dinero, personal y tiempo son esenciales para un esfuerzo de M&E completo. Si no se dispone de personal capacitado para efectuar una evaluación—o no pueden sacar al personal de otras funciones para que hagan un M&E sólido—y si no se puede pagar a un evaluador externo, entonces se puede realizar sólo una evaluación básica del progreso del programa.

- Si se dispone de pocos recursos, la primera prioridad debe ser establecer un sistema de monitoreo

Si los recursos son muy limitados, su mejor uso sería establecer un sistema de monitoreo efectivo, para asegurar y documentar que el programa se implementó de acuerdo al plan.

- Si se dispone de recursos adicionales, hacer algún tipo de evaluación de proceso

Algunos tipos de evaluaciones de proceso se pueden hacer a bajo costo, por ejemplo, con supervisores que periódicamente observen la entrega de servicios o entrevisten a los clientes del programa como parte de sus funciones. Las evaluaciones de proceso más sistemáticas (tales como efectuar grupos focales con los jóvenes) requieren más recursos.

- Las evaluaciones de resultados requieren un nivel de recursos cuyos costos son entre moderados y altos

Se necesita decidir al inicio si se hará una evaluación de resultados para presupuestarla como es debido. Los costos en gran parte dependerán de cuántos resultados se quiera medir y el nivel de dificultad de la medición. También dependerá de los datos que ya existan y cuánta información nueva se quiera recolectar. Los siguientes pasos pueden ayudar a controlar los costos de las evaluaciones de resultados:

- Limitar los resultados que se examinarán sólo a los que son más importantes para el programa.
- Elegir resultados que pueden medirse usando los métodos de recolección de datos menos costosos.<sup>26</sup>

<sup>26</sup> Los métodos de recolección de datos se tratarán en el Capítulo 4.

- *Elegir indicadores para los cuales ya existe información.*
- **Las evaluaciones de impacto requieren un nivel aun mayor de recursos financieros y técnicos**

Las evaluaciones de impacto sólo deben efectuarse cuando existe una razón importante para hacerlo. Tales razones pueden ser demostrar la eficacia de una estrategia del programa en una población objetivo particular, o satisfacer requisitos gubernamentales o de los donantes.

- **Cualesquiera sean los montos de recursos disponibles, hay que usarlos creativamente**

Existen muchas formas de recolectar datos. Los programas a menudo recolectan demasiados datos. A veces recolectan datos sobre demasiados temas, o datos que no tienen relación con los objetivos o actividades, o usan diferentes métodos para recolectar los mismos datos de la misma población objetivo. Gastar muchos recursos en la recolección de datos no garantiza que se obtendrán resultados que ayuden a entender mejor el programa o a los participantes.

Administrar el esfuerzo de M&E requiere planificación y creatividad. Hay que pensar cuidadosamente sobre la información que se quiere recolectar. Hay que encontrar formas de recolectar datos que se relacionen con los resultados que se esperan lograr, el significado del programa para los participantes y los factores que influyen en el éxito del programa.

- **Presupuestar para un esfuerzo de M&E es una parte importante de la planificación**

La hoja de trabajo que aparece a continuación ayudará a pensar cómo calcular los costos de cada categoría en un presupuesto de M&E, y puede ser usada como referencia cuando se preparen estimaciones detalladas para cada rubro del presupuesto. Sin embargo, algunas decisiones—tales como qué indicadores y métodos de recolección de datos se utilizarán, y la frecuencia y coordinación de la recolección de datos—se basarán en la materia tratada en los Capítulos 4 al 8, y deberían ser consideradas antes de finalizar el presupuesto de M&E.

| <b>Hoja de trabajo 3.2</b>   |                          |                                    |                                    |
|--|--------------------------|------------------------------------|------------------------------------|
| <i>Preparación de un presupuesto de M&amp;E</i>  |                          |                                    |                                    |
| <b>Rubro</b>   | <b>Fondos necesarios</b> | <b>Fuente del apoyo financiero</b> | <b>Fuente del apoyo en especie</b> |
| <b>Salarios:</b> para el personal que se necesita en asistencia técnica, recolección de datos, ingreso de datos y análisis (personal, entrevistadores, supervisores, chóferes, etc). |                          |                                    |                                    |
| <b>Viáticos:</b> costos diarios de alojamiento y alimentación.   |                          |                                    |                                    |
| <b>Viajes:</b> tarifas de buses y taxis, gasolina, alquiler de vehículos y mantenimiento.  |                          |                                    |                                    |
| <b>Imprenta:</b> cuestionarios de la encuesta, guías de entrevistas, informes finales, etc.  |                          |                                    |                                    |

|  |  |  |  |
|--|--|--|--|
| <b>Equipo:</b> bicicletas, computadoras y mantenimiento.   |  |  |  |
| <b>Comunicaciones:</b> teléfonos, fax, computadoras, radios, franqueo postal.  |  |  |  |
| <b>Suministros:</b> papel, disquetes de computación, lápices, carpetas, etc.   |  |  |  |
| <b>Actividades de diseminación:</b> costos de seminarios o conferencias, refrescos, materiales, carpetas, materiales de presentación, etc. |  |  |  |
| <b>TOTAL</b>   |  |  |  |

### ¿Quién debería estar involucrado en el monitoreo y la evaluación?

- Los esfuerzos de M&E deberían involucrar a muchos interesados

Muchas personas en la comunidad están interesadas en el M&E. La lista de interesados puede incluir a personal del programa, jóvenes, administradores y profesores de las escuelas, padres, líderes de la comunidad, funcionarios del gobierno local, proveedores de servicios y donantes. Estos pueden querer participar activamente o involucrarse en alguna o todas las fases de una evaluación: planificación y diseño, recolección y análisis de datos, identificación de los hallazgos, conclusiones y recomendaciones claves de una evaluación, diseminación de los resultados y, finalmente, planificación de cómo los resultados de una evaluación se pueden utilizar para mejorar un programa.

- La participación de los interesados puede hacer que los esfuerzos de M&E sean más relevantes y efectivos

La evaluación participativa facilita la identificación de las necesidades y prioridades locales y pone los temas de evaluación en el contexto de la vida de las personas. Un proceso participativo puede ayudar a lograr las siguientes metas relacionadas al M&E:

- *Desarrollar consenso acerca de los temas clave a ser abordados en una evaluación;*
- *Identificar qué información necesitan los interesados en el programa;*
- *Asegurar que el personal del programa entienda la necesidad de evaluar su papel en la implementación y cómo se usarán los resultados para mejorar el programa;*
- *Evitar métodos de evaluación intrusivos o inadecuados, por ejemplo, en cómo se recoge la información y*
- *Mantener abiertas las líneas de comunicación entre los interesados para facilitar posteriormente la diseminación y debate de los resultados de la evaluación.*

Los interesados también pueden ayudar a informar a evaluadores externos acerca del contexto del programa, y pueden facilitar un contacto continuo entre los que efectúan la evaluación y los que se ven afectados por ella.<sup>27</sup>

<sup>27</sup> Lawrence, 1989.

- La evaluación participativa es una forma de involucrar a los interesados más importantes—los jóvenes

Los jóvenes que son el objetivo del programa son los interesados más importantes. Sin embargo, algunos adultos administradores de programas y también el personal pueden encontrar dificultades en trabajar con jóvenes regularmente, debido a las muchas diferencias que a veces existen entre generaciones en cuanto a actitudes, comportamientos y creencias.

La *evaluación participativa* es un conjunto de técnicas que enfatizan la participación de la comunidad en la acumulación de conocimiento, y que ayudan a las personas involucradas a entender mejor sus áreas de preocupación o inquietudes. Este conocimiento basado en la experiencia ayuda a dirigir el desarrollo del programa y define la gama de servicios ofrecidos. La participación generalmente se da a través de todas las fases de la evaluación: planificación y desarrollo, recolección y análisis de datos, diseminación de los resultados y preparación de un plan para mejorar el desempeño del programa.<sup>28</sup> Los planificadores de programas en los Estados Unidos han encontrado algunas estrategias efectivas para trabajar con jóvenes, los cuales se pueden aplicar en muchos ambientes sociales. Éstas se presentan en el siguiente cuadro.

| <b>Consejos para involucrar a los jóvenes en M&amp;E participativa<sup>29</sup></b>  |  |
|--|--|
| <b>Consejo</b>   | <b>Ejemplos</b>  |
| <ul style="list-style-type: none"> <li>• Integrar a los jóvenes en los esfuerzos de los programas y en la planificación de M&amp;E.</li> </ul>   | <ul style="list-style-type: none"> <li>• Organizar las reuniones en lugares accesibles. Mantener la comunicación y transmitir la información necesaria. Fomentar la participación total y el derecho a voto.</li> </ul>                      |
| <ul style="list-style-type: none"> <li>• Ser abierto y no tener prejuicios acerca de los pensamientos y sugerencias de los jóvenes.</li> </ul>   | <ul style="list-style-type: none"> <li>• Tener cuidado de no descartar o reaccionar negativamente a las sugerencias de los jóvenes. Darles tiempo para sentirse cómodos y participar plenamente. Solicitar sus ideas y opiniones.</li> </ul> |
| <ul style="list-style-type: none"> <li>• Sacar provecho de la experiencia que ofrecen los jóvenes.</li> </ul>  | <ul style="list-style-type: none"> <li>• Los jóvenes saben mucho acerca de sus pares. Hay que estimularlos a compartir sus perspectivas acerca de los efectos positivos o negativos de los programas.</li> </ul>                             |
| <ul style="list-style-type: none"> <li>• Ser honesto acerca de las expectativas de los programas, las contribuciones de los jóvenes y los beneficios de la participación juvenil.</li> </ul> | <ul style="list-style-type: none"> <li>• No pretender que el programa puede resolver todos los problemas. Ser realista acerca de lo que se puede emprender.</li> </ul>   |
| <ul style="list-style-type: none"> <li>• Estar preparado para ofrecer apoyo a los jóvenes.</li> </ul>  | <ul style="list-style-type: none"> <li>• El apoyo puede incluir consejería, asistencia financiera, transporte, capacitación, supervisión e información.</li> </ul>   |
| <ul style="list-style-type: none"> <li>• Hacer que el trabajo sea interactivo y entretenido</li> </ul>   | <ul style="list-style-type: none"> <li>• Los jóvenes tienen más probabilidades de interesarse y permanecer interesados en programas que les resulten informativos, entretenidos y que les den satisfacción.</li> </ul>                       |
| <ul style="list-style-type: none"> <li>• Ayudar a aumentar las habilidades de los jóvenes para que puedan involucrarse más.</li> </ul>   | <ul style="list-style-type: none"> <li>• Los jóvenes pueden necesitar información y ayuda para sentirse cómodos a medida que comienzan a comunicarse con distintas audiencias.</li> </ul>  |

<sup>28</sup> USAID CDIE, 1996

<sup>29</sup> Adaptado de Clark, Haughton-Denniston, Flinn, et.al., 1993, citado en Brindis y Davis, 1998: volumen 4, página 49.

- **Involucrar a los interesados y los jóvenes puede traer problemas**

Las desventajas de involucrar a los interesados, especialmente de otras organizaciones, incluyen las siguientes:

- *Puede resultar difícil ser objetivo en la selección de jóvenes y organizaciones representativas para participar en la evaluación.*
- *Los interesados pueden carecer de conocimiento acerca de cómo funciona un programa.*
- *Las organizaciones pueden tener percepciones competitivas y preocupaciones que sean difíciles de resolver o clasificar según su prioridad.*
- *La capacidad de realizar una evaluación independiente puede verse comprometida al incluir a diversas organizaciones.*
- *Un mayor número de participantes puede requerir una mayor asignación del tiempo del personal y los recursos del programa.<sup>30</sup>*

### **¿Quién debería efectuar la evaluación?**

Las evaluaciones pueden efectuarlas el personal del programa, personas fuera del programa o una combinación de los dos. Al decidir quién efectuará una evaluación, deben considerarse varios temas. En primer lugar, ¿cuál es la estructura más apropiada para el equipo de evaluación? En segundo lugar, ¿qué es factible? ¿Cuánto se puede gastar, dado los recursos presupuestarios? Puede resultar demasiado complicado o ineficiente involucrar a todos los interesados en cada actividad de M&E.

- **Usar al personal del programa para efectuar las evaluaciones tiene ventajas**

La ventaja de usar personal interno es que éste ya está familiarizado con el programa y puede ser capacitado rápidamente. También los integrantes del personal pueden estar conscientes de las fortalezas y debilidades particulares de los programas y de aquellos que requieren atención. Finalmente, los resultados de la evaluación serán de máxima utilidad al personal del programa, el cual está en condiciones de modificar y mejorarlo según sea necesario. Usar al personal propio también puede ser más factible financieramente, ya que los evaluadores externos a menudo son más costosos. Además, por razones financieras o logísticas, los evaluadores externos pueden estar disponibles sólo por un tiempo limitado.

- **Usar evaluadores externos es más apropiado en algunas situaciones**

A pesar de las ventajas de usar personal propio, puede haber situaciones cuando es más apropiado usar evaluadores externos o involucrar a evaluadores externos con el personal del programa. Las agencias financieras a veces requieren que por lo menos una parte de las evaluaciones sean efectuadas por evaluadores externos. Como ellos tienen menos interés en el resultado de la evaluación, los evaluadores externos son percibidos más objetivamente y, por ende, sus resultados suelen tener más credibilidad. Sin embargo, a pesar de que deben mantener su objetividad, los evaluadores externos deben ser sensibles a las metas de los programas y al contexto local en el que se está implementando el programa. En vez de

---

<sup>30</sup> Adaptado de Clark, Houghton-Denniston, Flinn, y cols., 1993, citado en Brindis y Davis, 1998: volumen 4, página. 49.

representar una amenaza, los evaluadores deben ser considerados parte del sistema de apoyo del programa.

- **Cuando los recursos de personal son limitados, puede resultar más factible usar evaluadores externos**

Usar evaluadores internos o externos también depende de la disponibilidad de tiempo y experiencia del personal del programa, ya que las evaluaciones pueden ser muy exigentes. Se necesitará evaluar la experiencia y habilidades del personal para efectuar el M&E, y cuánto tiempo tendrán que invertir en estos esfuerzos. También se necesita considerar qué personal debe involucrarse, cómo sus vacaciones y feriados afectan su disponibilidad, y si se necesita ayuda externa. Idealmente, los jóvenes y otros interesados importantes deberían participar hasta donde sea posible. En algunos casos, la evaluación externa puede incluir asistencia técnica que, como parte de un enfoque amplio, puede capacitar al personal interno y aumentar la efectividad del programa.

### **¿Dónde debería efectuarse el monitoreo y la evaluación?**

Si el programa funciona en uno o dos lugares o cubre un área geográfica pequeña, se pueden efectuar los esfuerzos de monitoreo y evaluación más fácilmente abarcando toda el área, o conjunto de lugares. Sin embargo, si el programa cubre un área más grande o tiene varias ubicaciones, puede ser necesario reducir el alcance geográfico del esfuerzo. Cómo seleccionar las ubicaciones o áreas que se incluirán en el esfuerzo de M&E depende de las necesidades de información y de los recursos financieros y humanos.

- **Hacer un esfuerzo para monitorear todas las ubicaciones del programa**

Como el monitoreo es esencial para la administración efectiva del programa, se debe tratar de incluir todas las ubicaciones del programa en la recolección de información básica— como por ejemplo, si las actividades planeadas se han completado y se ha identificado el número y el género de los clientes que han sido servidos por el programa. Esto proporcionará una visión de cómo progresa la implementación del programa y también permite comparar cómo se desempeñan las distintas ubicaciones. Si algunas de las ubicaciones del programa tienen una mayor capacidad para recolectar datos que otras, se puede considerar hacerlas recolectar datos de monitoreo adicionales que ayudarán a contestar otras preguntas acerca de la implementación del programa.

Si no es posible recolectar los mismos datos de monitoreo en todas las ubicaciones, quizás sea mejor no desarrollar el programa en ese lugar, a no ser que exista la absoluta seguridad de que la estrategia puede funcionar sin monitoreo. Para muchas estrategias, por ejemplo, la educación de pares, el monitoreo es esencial para asegurarse de que el programa está siendo implementado tal como fue planeado. Si un programa puede funcionar sin monitoreo, y se decide implementarlo, entonces sería recomendable que al menos se recogieran datos de una muestra de sitios considerados “representativos” de todo el programa. Cómo elegir una muestra representativa de ubicaciones se trata en el Capítulo 6.

- **Probablemente se necesitará limitar la recolección de datos para la evaluación**

Efectuar evaluaciones usualmente requiere que se limite la recolección de datos en ciertas formas. Por ejemplo, en una evaluación de procesos rara vez es posible evaluar cada contacto de servicio u obtener retroalimentación de todos los participantes del programa. En programas más grandes, también se tendrán que limitar las evaluaciones de procesos a sólo una muestra de las ubicaciones del programa.

La mayoría de las evaluaciones de resultados e impactos también requieren que se aplique algunas restricciones sobre dónde se recolectarán los datos. Esto es especialmente cierto cuando los objetivos del programa están relacionados a los resultados medidos para la población general de jóvenes. Aquí, usualmente será necesario recolectar datos de jóvenes en un subconjunto o una muestra de las áreas geográficas cubiertas por el programa.

- **Elegir lugares para efectuar evaluaciones requiere cuidadosa consideración**

A veces un programa no tiene un área geográfica de influencia claramente definida. Si el área de influencia de un programa está definida con mucha amplitud, tal como una ciudad completa o una región, entonces puede ser más difícil medir los cambios en los objetivos, aún si el programa funciona en forma óptima.

Para elegir ubicaciones o áreas geográficas en las cuales efectuar evaluaciones, hay que hacerse las siguientes preguntas clave:

- *¿A qué área geográfica llega el programa?*
- *¿En cuántas ubicaciones o áreas geográficas es necesario efectuar una evaluación fuerte?*
- *¿Estas ubicaciones elegidas, representan adecuadamente las características de la población objetivo juvenil y del programa que se está implementando?*
- *¿Cuántas observaciones se necesitan por ubicación o área geográfica?*
- *¿Cómo se deben elegir las ubicaciones o áreas geográficas?*

Muchos de estos temas serán abordados en los Capítulos 5 y 6. Las siguientes consideraciones prácticas probablemente también influirán sobre la decisión en cuanto adónde evaluar:

- *¿Existen datos disponibles en todas las ubicaciones?*
- *¿Cuán fácil es recolectar nuevos datos en cada ubicación?*
- *¿Cómo afectará la recolección de datos el desempeño de las actividades normales del programa?*
- *¿Existen otros recursos disponibles para ayudar a recolectar y analizar los datos? (Por ejemplo, universidades locales o grupos de investigación)*

### ***...¿Qué preguntas deben hacerse para determinar el tipo de esfuerzo de M&E?***

La siguiente lista de verificación y flujograma se pueden usar para determinar el tipo de actividades de M&E que se podrían realizar. Estas preguntas deben responderse antes de usar el flujograma:

- *¿Están claras las metas del programa?*
- *¿Están relacionados los objetivos a las metas y a los resultados esperados?*
- *¿Están los objetivos expresados en términos medibles?*
- *¿Están definidas las actividades?*

- ¿Están relacionadas las actividades a los objetivos del programa?

Este flujograma se utiliza para determinar qué tipo de esfuerzo de M&E se puede realizar. Cuando se llegue a un cuadro en negrilla, el contenido identifica el tipo de M&E que es mas apropiado para su programa.

|   |                             |   |                                  |                                  |                          |                                      |
|---|-----------------------------|---|----------------------------------|----------------------------------|--------------------------|--------------------------------------|
| ¿Se sabe si la estrategia del programa funciona?  | Sí<br>M, P                  |   |                                  |                                  | No<br>M, P, O, I         |                                      |
| ¿Se sabe si la estrategia del programa funciona en la población objetivo a la que se quiere llegar? | <b>Sí</b><br><b>M, P</b>    |   |                                  |                                  | No<br>M, P, O, I         |                                      |
| ¿El proyecto está al comienzo, a la mitad o en el final?  |                             |   |                                  | Comienzo<br>M, P, O, I           | Mitad<br>M, P, O         | Final<br>P, O                        |
| ¿Qué nivel de recursos financieros tiene?   | <b>Bajos</b><br><b>M, P</b> | Medianos<br>M, P, O                     | Altos<br>M, P, O, I              | Medianos o altos<br>M, P, O      | <b>Bajos</b><br><b>P</b> | Medianos o altos<br>P, O             |
| ¿Qué nivel de recursos de personal tiene?   | <b>Bajo</b><br><b>M, P</b>  | <b>Mediano o alto</b><br><b>M, P, O</b> | <b>Mediano</b><br><b>M, P, O</b> | <b>Alto</b><br><b>M, P, O, I</b> | <b>Bajo</b><br><b>P</b>  | <b>Mediano o alto</b><br><b>P, O</b> |

- M = Efectuar monitoreo
- P = Efectuar evaluación de procesos
- O = Efectuar evaluación de resultados
- I = Efectuar evaluación de impacto

### ***...¿Qué se necesita para llevar a cabo cada tipo de evaluación? (Cómo usar el resto de este manual)***

El cuadro que aparece a continuación ayudará a determinar cómo utilizar el resto de este manual. Dependiendo del alcance del esfuerzo de M&E, el cuadro guía hacia los indicadores relevantes de SSRA (Capítulo 4 y Cuadros de Indicadores), fuentes de datos y métodos de recolección de datos (Capítulo 4), y a los instrumentos correspondientes (Sección II de este manual). Para las evaluaciones de resultados e impactos, se sugieren diseños de estudios en el Capítulo 5. El cuadro que sigue apunta a temas relevantes de muestreo, que se explican en el Capítulo 6, y a los tipos de análisis, descritos en el Capítulo 7.



## Cómo usar el resto de este manual

|  | Nivel de recursos necesitados | Selección de indicadores (Cap. 4)  | Fuentes de datos, métodos y recolección (Cap. 7)   | Instrumentos y herramientas de recolección de datos (Sección II de este manual)  | Diseño de estudios (Cap. 5) | Muestreo (Cap. 6) | Recolección de datos y análisis (Cap. 7 y 8)   |
|--|-------------------------------|--|--|--|-----------------------------|-------------------|--|
| <b>M<br/>o<br/>n<br/>i<br/>t<br/>o<br/>r<br/>e<br/>o</b> | Bajo                          | Indicador Cuadro 2: Desarrollo de sistemas; Indicador Cuadro 3: Implementación | Estadísticas de servicios; informes y documentos administrativos; registro de eventos; otros tipos de registros; encuestas | Instrumento 2: Hojas de recuento; Instrumento 3: Formularios de informes; Instrumento 5: Índices compuestos ilustrativos; Instrumento 6: Inventario de instalaciones y servicios | N/D                         | Uso limitado      | Comparar desarrollo de sistemas y funcionamiento de indicadores de implementación con los objetivos; comparar indicadores de distintas ubicaciones del programa; evaluar tendencias de los indicadores a lo largo del tiempo |

## Cómo usar el resto de este manual

|  | Nivel de recursos necesitados | Selección de indicadores (Cap. 4)  | Fuentes de datos, métodos, recolección (Cap. 4)   | Instrumentos y herramientas de recolección de datos (Sección II de este manual)  | Diseño de estudios (Cap. 5) | Muestreo (Cap. 6)  | Recolección de datos y análisis (Cap. 7 y 8)   |
|--|-------------------------------|--|---|--|-----------------------------|--|--|
| <b>P<br/>r<br/>o<br/>c<br/>e<br/>s<br/>o<br/>s</b>             | Bajo a moderado               | Indicador Cuadro 1: Diseño; Indicador Cuadro 2: Desarrollo de Sistemas; Indicador Cuadro 3: Implementación | Informes y documentos administrativos; registro de eventos; otros tipos de registros; revisión de documentos; visitas a las localidades; observaciones directas; entrevistas con informantes claves (proveedores de servicios, administradores); retroalimentación no estructurada de los clientes; entrevistas finales a los clientes; clientes simulados; grupos focales o sesiones informales de oyentes | Instrumento 1: Listas de Referencia; Instrumento 2: Hojas de recuento; Instrumento 6: Inventario de Instalaciones y Servicios; Instrumento 7: Manual de Observación para Orientación y Procedimientos Clínicos; Instrumento 8: Manual de Entrevistas para Personal que Proporciona Servicios de SR; Instrumento 9: Manual para Entrevistas Finales a Clientes; Instrumento 10: Cuestionario para Rendición de Cuentas de Clientes simulados; Instrumento 14: Evaluando la Hoja de Trabajo de Efectividad de la Coalición | N/D                         | Elegir muestras de ubicaciones del programa, participantes, proveedores de servicios, transacciones de servicios, etc., para medir el desarrollo/ funcionamiento de los sistemas y los indicadores de implementación | Comparar el desarrollo y funcionamiento de los sistemas y los indicadores de implementación con los objetivos o estándares; comparar indicadores para diferentes ubicaciones del programa; evaluar tendencias en los indicadores a lo largo del tiempo |
| <b>R<br/>e<br/>s<br/>u<br/>l<br/>t<br/>a<br/>d<br/>o<br/>s</b> | Moderado                      | Indicador Cuadro 4: Resultados   | Estadísticas de servicios; encuestas; encuestas de población; grupos focales o sesiones informales de oyentes   | Instrumento 3: Formularios de Informes; Instrumento 6: Inventario de Instalaciones y Servicios; Instrumento 11: Cuestionario para la Comunidad; Instrumento 12: Encuesta Global; Instrumento 13: Manual de Discusión para el Grupo Focal   | N/D                         | Elegir muestras de los jóvenes, localidades para el programa, comunidades, etc. para la medición de los indicadores de resultados  | Comparar los indicadores de resultados con los objetivos; comparar indicadores para las diferentes ubicaciones del programa; evaluar las tendencias de los indicadores   |

| Cómo usar el resto de este manual |                              |                                   |   |  |                             |   |   |
|-----------------------------------|------------------------------|-----------------------------------|---|--|-----------------------------|---|---|
|                                   | Nivel de recursos necesarios | Selección de indicadores (Cap. 4) | Fuentes de datos, métodos, recolección (Cap. 4)   | Instrumentos y herramientas de recolección de datos (Sección II de este manual)  | Diseño de estudios (Cap. 5) | Muestreo (Cap. 6)   | Recolección de datos y análisis (Cap. 7 y 8)                                    |
| I<br>m<br>p<br>a<br>c<br>t<br>o   | Alto                         | Indicador Cuadro 4: Resultados    | Estadísticas de servicios; encuestas; encuestas de población; grupos focales o sesiones informales de oyentes | Instrumento 3: Formularios de Informes; Instrumento 6: Inventario de Instalaciones y Servicios; Instrumento 11: Cuestionario para la Comunidad; Instrumento 12: Encuesta Global; Instrumento 13; Manual de Debate para el Grupo Focal. | Capítulo 5                  | Escoger muestras de jóvenes, localidades para programas, comunidades, etc., en cada grupo experimental o de comparación para la medición de indicadores de resultados | Comparar indicadores de resultados para grupos de “tratamiento” y de “control”. |

## □ Capítulo 4: Indicadores

### Una mirada al capítulo:

- *Define y explica los indicadores*
- *Proporciona ejemplos de cómo seleccionar y modificar los indicadores para que coincidan con los objetivos y actividades del programa*

### ...¿Qué es un indicador?

Un *indicador* es una declaración medible de los objetivos y actividades del programa. Una vez que se han definido los objetivos y actividades del programa, se pueden desarrollar indicadores—o medidas—para cada objetivo y actividad. Algunos programas pueden tener un indicador único; otros tienen múltiples indicadores. Generalmente es preferible tener varios indicadores para capturar las múltiples dimensiones del programa. Sin embargo, se debe seleccionar cuidadosamente un número manejable de indicadores que reflejen verazmente los objetivos y actividades del programa, y las prioridades de la evaluación.

Continuando con el ejemplo presentado en el Capítulo 3, el siguiente cuadro muestra qué tipos de indicadores se pueden usar para medir los objetivos y actividades asociados con el retraso de la edad de iniciación sexual, a través de programas de educación de pares.

| <b>Determinación de objetivos, indicadores y actividades</b>   |  |  |  |
|--|--|--|--|
| <b>Objetivo</b>  | <b>Posibles indicadores de resultados</b>  | <b>Actividades</b>   | <b>Posibles indicadores de programas</b>   |
| Objetivos a nivel de población:  |  |  |  |
| <ul style="list-style-type: none"> <li>Atrasar la edad de la iniciación sexual en jóvenes de 14 a 19 años</li> </ul>   | <ul style="list-style-type: none"> <li>Edad promedio de la iniciación sexual entre jóvenes de 14 a 19 años</li> </ul>  | <ul style="list-style-type: none"> <li>Promover la disponibilidad de servicios de orientación de pares a través de anuncios en la radio</li> </ul> | <ul style="list-style-type: none"> <li>Número de avisos radiales que salen en el aire</li> </ul>   |
| <ul style="list-style-type: none"> <li>Aumentar a 25% el porcentaje de jóvenes entre 14 y 19 años que buscan servicios de orientación brindados por adolescentes educadores</li> </ul> | <ul style="list-style-type: none"> <li>Porcentaje de jóvenes entre 14 y 19 años que buscan servicios de orientación brindados por adolescentes educadores</li> </ul> | <ul style="list-style-type: none"> <li>Implementar un "rincón de orientación de pares" en cinco clínicas de salud</li> </ul>                       | <ul style="list-style-type: none"> <li>Número de nuevos avisos radiales que salen en el aire</li> </ul>  |
|  |  | <ul style="list-style-type: none"> <li>Dar charlas informativas de adolescentes educadores en las escuelas dos veces a la semana</li> </ul>        | <ul style="list-style-type: none"> <li>Número de clínicas que tienen rincones de orientación de adolescentes</li> </ul>  |
|  |  | <ul style="list-style-type: none"> <li>Proporcionar servicios de orientación de calidad por adolescentes educadores</li> </ul>                     | <ul style="list-style-type: none"> <li>Número de días por semana que el rincón de orientación tiene personal disponible</li> </ul>                                   |
|  |  |  | <ul style="list-style-type: none"> <li>Número de clientes que buscan servicios de orientación</li> </ul>   |
|  |  |  | <ul style="list-style-type: none"> <li>Número de charlas informativas de adolescentes educadores en las escuelas</li> </ul>  |
|  |  |  | <ul style="list-style-type: none"> <li>Número de jóvenes que asisten a las charlas informativas</li> </ul>   |
|  |  |  | <ul style="list-style-type: none"> <li>Calidad de las presentaciones de los adolescentes educadores (basado en los criterios para los planes de estudio)</li> </ul>  |
|  |  |  | <ul style="list-style-type: none"> <li>Proporción de clientes que consideran la orientación de alta calidad en las entrevistas de salida</li> </ul>                  |
|  |  |  | <ul style="list-style-type: none"> <li>Puntuación asignada a los adolescentes orientadores por los observadores (basados en los criterios de orientación)</li> </ul> |

| Objetivo   | Posibles indicadores de resultados  | Actividades   | Posibles indicadores de programas   |
|--|---|---|---|
| Objetivos a nivel de programa:<br><ul style="list-style-type: none"> <li>Aumentar la capacidad de los adolescentes educadores para proporcionar orientación a los jóvenes</li> </ul> | <ul style="list-style-type: none"> <li>Número de adolescentes educadores que están calificados para proporcionar orientación a los jóvenes</li> </ul> | <ul style="list-style-type: none"> <li>Reclutar adolescentes educadores de un grupo de adolescentes que asisten a la clínica</li> </ul> | <ul style="list-style-type: none"> <li>¿Se ha completado el reclutamiento? (Sí/No)</li> </ul>   |
|  |   | <ul style="list-style-type: none"> <li>Seleccionar a 30 adolescentes educadores</li> </ul>  | <ul style="list-style-type: none"> <li>Número de adolescentes educadores seleccionados</li> </ul>   |
|  |   | <ul style="list-style-type: none"> <li>Desarrollar un plan de capacitación</li> </ul>   | <ul style="list-style-type: none"> <li>¿Se ha desarrollado un plan de capacitación? (Sí/No)</li> </ul>  |
|  |   | <ul style="list-style-type: none"> <li>Capacitar a los adolescentes educadores para proporcionar orientación</li> </ul>                 | <ul style="list-style-type: none"> <li>Número de “temas clave” cubiertos por el plan de capacitación comparado con la lista de verificación</li> </ul>                |
|  |   |   | <ul style="list-style-type: none"> <li>Proporción de adolescentes educadores que demuestran habilidades efectivas de orientación durante los juegos de rol</li> </ul> |

### Los indicadores se pueden expresar de diferentes formas

Como se puede ver en el ejemplo anterior, los indicadores se pueden expresar de distintas maneras. Los indicadores numéricos se expresan en conteos, porcentajes, proporciones, tasas o promedios. Los siguientes indicadores son conteos:

- Número de avisos radiales que salen al aire
- Número de clientes que buscan servicios de orientación de pares

Para efectos de evaluación, generalmente es más informativo dar los indicadores como porcentajes, coeficientes y proporciones. Estas medidas permiten ver lo que se logró en relación al *denominador*, o número total posible, mientras los conteos simplemente dan una idea del número de eventos que ocurrieron, o el número de personas alcanzadas, sin indicar el número total posible. Por ejemplo, se puede contar el número de jóvenes que han retrasado la edad de la iniciación sexual, pero si se tiene un *denominador*, el número total de jóvenes en un área geográfica determinada, se podrá calcular la proporción de jóvenes en esa área que han retrasado su iniciación sexual. Esto permitirá medir la cobertura del programa y los efectos sobre el comportamiento a nivel de la población.

Más adelante, en la Sección II de este manual, se proporciona una definición para cada uno de estos términos y se señala cómo calcular los diferentes tipos de indicadores numéricos.

Los indicadores no numéricos se expresan en palabras. También se les llama cualitativos o categóricos. Estos indicadores usualmente indican la presencia o ausencia de un evento o criterio, tal como:

- ¿Se ha completado el reclutamiento de capacitadores educadores? (Sí/No)

- *¿Los planes de capacitación incluyen temas relacionados a relaciones y sexualidad? (Sí/No)*

Los indicadores no numéricos también se pueden usar para resumir descripciones o evaluar calidad o cobertura. Esto se puede hacer creando un índice que consiste de una lista de asuntos a los que se les asigna un número a cada uno, que luego se suman para obtener una puntuación. Por ejemplo, en el cuadro que sigue, a cada uno de los puntos incluidos en la lista de verificación en la columna derecha se le asigna un punto, y luego se suman esos puntos para determinar la puntuación total de la presentación.

| <b>Uso de indicadores no numéricos para medir calidad</b> |   |
|---|---|
| <b>Indicador</b>  | <b>Índice y puntuación de calidad</b>   |
| La presentación del adolescente educador es integral      | Observar la presentación. Marcar cada tópico tratado adecuadamente. Darle un punto a cada ítem marcado y obtener el total para determinar la puntuación de calidad.   |
|   | <ul style="list-style-type: none"> <li>• Anatomía y reproducción</li> <li>• Abstinencia</li> <li>• Anticoncepción</li> <li>• Demostrar el uso de preservativos</li> <li>• Tomar la decisión de tener sexo</li> <li>• Cómo decir “no” al sexo</li> <li>• Resistir la presión de los pares para tener relaciones sexuales</li> <li>• Dónde buscar orientación</li> <li>• Dónde buscar servicios de salud</li> </ul> |
|   | <b>TOTAL</b>  |

### **Tal como los objetivos, los indicadores deben ser específicos**

Mientras más específico el indicador, mayor es la probabilidad que se puedan medir los objetivos y actividades con precisión. Los indicadores deben especificar:

- *las características de la población objetivo, tales como género, edad, residencia, estado civil y situación escolar,*
- *la ubicación de la población objetivo, tales como juventud urbana o rural, juventud en una determinada ciudad o distrito, juventud que participa en el programa o juventud que asiste a determinadas escuelas o clínicas, y*
- *el plazo en el cual se pretende lograr los objetivos.*

### **Un indicador debe tener la misma escala que el objetivo al que corresponde**

Los indicadores deben tener la misma escala que los objetivos del programa. Por ejemplo, si el objetivo es demorar la edad promedio de la iniciación sexual entre los jóvenes de 14 a 19 años que viven en el distrito donde se llevará a cabo el programa, el indicador debería medir la “edad promedio al momento de la iniciación sexual entre los jóvenes de 14 a 19 años que viven en el distrito X”. Si el indicador tiene una escala distinta al objetivo, los resultados serán engañosos.

## *...Tipos de indicadores*

### **Desarrollar indicadores para cada componente del programa que se va a medir**

Una vez que se ha definido el alcance del esfuerzo de M&E, se deben desarrollar diferentes indicadores para cada componente del programa que se va a medir. Por ejemplo, si se planea efectuar una evaluación de proceso, se deben desarrollar indicadores de diseño, desarrollo y funcionamiento de sistemas o implementación. Si se planea efectuar una evaluación de impacto, se deben desarrollar indicadores para medir la implementación y resultados del programa.

En este manual hemos categorizado los indicadores basados en los componentes del programa que serán monitoreados y evaluados. El Capítulo 10 contiene cuatro cuadros, cada uno con ejemplos de indicadores basados en aspectos del programa. Estos cuadros se pueden usar para seleccionar y adaptar los indicadores para que coincidan con el programa.

### **Los indicadores de diseño están relacionados a los “elementos clave”**

Los programas para jóvenes deberían ser diseñados en base a “elementos claves” de calidad. La experiencia internacional de programas para jóvenes, materno-infantiles, planificación familiar y VIH/SIDA, y la intuición y experiencia de los profesionales han producido varias recomendaciones clave para el diseño de programas para jóvenes. Algunos ejemplos son:

- *la existencia de metas y objetivos claramente definidos,*
- *los interesados locales están involucrados en la planificación del programa, y*
- *se ha realizado una evaluación de necesidades y se conocen las preferencias de la población objetivo de jóvenes hacia servicios de salud reproductiva.*

### **Los indicadores de desarrollo y funcionamiento de sistemas están relacionados a los objetivos y actividades programáticas**

Los objetivos programáticos indican los resultados en términos de la estructura organizacional, administración u operaciones de un programa, y las correspondientes actividades involucran el desarrollo y funcionamiento de los sistemas. Los indicadores de desarrollo y funcionamiento de los sistemas miden si los sistemas de una organización o de un programa están operando, y cuán efectivamente han preparado al personal del programa para la etapa de implementación. Ejemplos de indicadores de desarrollo y funcionamiento de sistemas incluyen:

- *número de adolescentes educadores capacitados para proporcionar orientación a jóvenes,*
- *existencia de una estructura organizacional clara, y*
- *número de asociaciones, redes o coaliciones establecidas para apoyar el programa de SSRA.*

### **Los indicadores de implementación están relacionados a los objetivos y actividades programáticas y de población**

Tanto los objetivos programáticos como los de población serán satisfechos con la implementación de las actividades del programa. Los indicadores de implementación miden si las actividades planificadas han sido efectuadas, cómo han sido efectuadas y la calidad de la implementación. Ejemplos de indicadores de implementación incluyen:

- *número de jóvenes que buscan servicios de orientación de pares,*
- *número de interesados y sus formas de participación en el programa de SSRA, y*
- *número y tipo de productos de comunicación desarrollados para la población objetivo.*

### **Los indicadores de resultados están relacionados a los objetivos a nivel de la población objetivo**

Los objetivos a nivel de la población indican los resultados en términos del participante en el programa, y son aseveraciones medibles de los resultados que se espera lograr en la población objetivo. Los indicadores de resultados miden el cambio en los resultados que las actividades del programa producen en la población objetivo. Ejemplos de indicadores de resultados incluyen:

- *edad promedio al momento de la iniciación sexual,*
- *porcentaje de jóvenes que dicen que ellos abogarían por comportamientos más saludables entre sus pares y amigos,*
- *tasa de embarazos entre las jóvenes durante un período específico de tiempo, y*
- *tasa de incidencia de ITS entre los jóvenes durante un período específico de tiempo.*

### ***...¿Cómo deberían ser enunciados los indicadores?***

Describir de forma clara y precisa los indicadores producirá resultados significativos en el esfuerzo de M&E.

### **Evaluar los indicadores en términos de su importancia y facilidad para recolectar datos**

Los indicadores se considerarían de alta prioridad si se puede aplicar uno o más de los siguientes criterios:

- *El indicador es prioritario dado el propósito y alcance de la evaluación*
- *El indicador experimenta con un enfoque nuevo*
- *El personal quiere tener información sobre el indicador*
- *Los jóvenes han identificado al indicador cómo importante*
- *Los donantes requieren información que el indicador medirá*



Si se determina que los datos que se necesitan para calcular los indicadores no están disponibles, se necesitará recoger nueva información. Es importante evaluar cuán fácil o difícil será la recolección de estos datos. Los factores a considerar para determinar la facilidad de la recolección de datos son:

- *Sensibilidad de los temas (especialmente en términos de las normas locales y el contexto cultural local)*
- *Recursos de personal y conocimientos especializados*
- *Requerimientos logísticos (transporte, impresión, vehículos)*
- *Tiempo*
- *Costo*
- *Jerga, modismos locales y términos profesionales usados para referirse al tema*

### **Formular los indicadores en un lenguaje claro y preciso**

Es importante usar palabras y frases claras y precisas para formular los indicadores. Los indicadores demasiado generales se prestan a muchas interpretaciones y pueden limitar la capacidad de interpretar los resultados del M&E.

Por ejemplo, un indicador general puede ser “el número de jóvenes que busca servicios de orientación de pares”. Este indicador debería ser formulado más precisamente como “número de jóvenes entre 14 y 19 años que residen en el distrito y que buscan servicios de orientación de adolescentes educadores durante un período de seis meses”.

### **Evitar los cambios de redacción en los indicadores después de comenzado el esfuerzo de M&E**

Cambiar la redacción de los indicadores durante la implementación del programa puede obstaculizar la capacidad de interpretar los resultados del M&E. Por ejemplo, se puede asumir que el indicador es “número de jóvenes entre 14 y 19 años que residen en el distrito y que buscan orientación de adolescentes educadores durante un período de seis meses”. Si a mitad de camino de la implementación del programa se cambia la especificación para contar el número de jóvenes entre 14 y 16 años que buscan servicios de orientación, puede aparecer que el número de clientes ha bajado. Por lo tanto, los resultados sugerirán que menos jóvenes están utilizando el programa, cuando en realidad esto puede no ser cierto.

Si ya se ha comenzado el esfuerzo de M&E y se descubre que los indicadores no son suficientemente específicos, es aconsejable *agregar* indicadores en vez de cambiar los existentes. Por ejemplo, si se encuentra que los jóvenes que buscan servicios de orientación en su mayoría tienen entre 12 y 15 años, se puede agregar el indicador “número de jóvenes entre 12 y 15 años que residen en el distrito y que buscan servicios de orientación de adolescentes educadores durante un período de seis meses”. Así se continuaría midiendo el indicador original para jóvenes entre 14 y 19 años, además del nuevo indicador para jóvenes entre 12 y 15 años.

### **Los indicadores deberían ser consistentes en el tiempo**

Los indicadores usados deberían ser consistentes durante toda la duración del esfuerzo de monitoreo y evaluación. Si se eliminan, agregan o modifican indicadores durante la implementación del programa, puede no ser posible evaluar el porqué de los cambios que están ocurriendo en la población objetivo. Por ejemplo, considérese los siguientes indicadores de ITS:

- *Porcentaje de jóvenes que reportan síntomas específicos de ITS*

Se supone que para medir este indicador inicialmente se desarrolló una lista de verificación de seis síntomas posibles que los adolescentes orientadores usan para registrar el diagnóstico, basado en lo que les informan sus clientes. Después de seis meses, se revisan los registros de las clínicas en cuatro clínicas del área de captación y se encuentra que los jóvenes han sido diagnosticados con ITS usando síntomas diferentes a los de la lista de verificación. Entonces se agregan otros cuatro síntomas a la lista de verificación que usan los adolescentes educadores. Esto significa que los adolescentes educadores pueden comenzar a registrar a jóvenes que mencionen cualquiera de estos cuatro síntomas adicionales, cuando antes no hubieran sido incluidos. Por lo tanto, si el porcentaje de jóvenes que reporta síntomas de ITS aumenta, no se sabrá si el cambio ocurrió por un verdadero aumento en la prevalencia de ITS o simplemente porque se agregaron cuatro criterios posibles adicionales a la lista de verificación.

### **Determinar cuidadosamente la dimensión temporal de los indicadores de resultados**

La mayor parte de los indicadores de resultados se refieren a los resultados esperados en el mediano o largo plazo. Por ejemplo, puede tomar varios años documentar cambios en los resultados en la tasa de embarazo entre las jóvenes. Por lo tanto, lo que se define como mediano y largo plazo variará de acuerdo a la naturaleza y complejidad de los objetivos y actividades del programa.<sup>31</sup> Por ejemplo, algunos programas pueden definir los resultados de mediano plazo como aquellos alcanzados dentro de un año, y los resultados de largo plazo como aquellos alcanzados en cinco años.

Debe establecerse un plazo razonable de tiempo para lograr los resultados deseados. Los programas para jóvenes a menudo están bajo presión para demostrar resultados y, por eso, tratan de medir cambios en un plazo corto, poco realista dado el tiempo y los objetivos. Los resultados entonces pueden indicar falsamente que no se han cumplido los objetivos. Una vez que se determina el plazo que se piensa tomará alcanzar los objetivos, se puede indicar el plazo de los indicadores de resultados. Luego será necesario darle seguimiento a los indicadores de resultados por un plazo suficientemente largo como para poder observar los cambios esperados.

### **Los indicadores deben ser válidos y confiables**

Los indicadores deben ser **válidos**, lo que significa que miden verazmente el concepto o evento que deben medir. También deben ser **confiables**, midiendo el asunto o evento consistentemente cada vez. Evaluar la validez y confiabilidad de los indicadores ayuda a asegurar que se minimicen los *errores* de medición.

Dos pasos pueden fortalecer la validez de los indicadores.

1. Desarrollar indicadores cuyo contenido incluya muestras adecuadas de todos los posibles significados de un concepto. Por ejemplo, para medir la calidad de la interacción entre los jóvenes y sus padres, se tiene que pensar en todos los posibles significados de calidad de la interacción. Se puede determinar cuán a menudo se comunican los jóvenes con sus padres, cuán largas son las conversaciones, qué temas se discuten y las percepciones del joven acerca de la interacción. Todos son aspectos que contribuyen a la calidad de la interacción. Se puede desarrollar entonces una serie de indicadores que juntos midan la calidad de la interacción, tal como:

---

<sup>31</sup> Muchos de los indicadores incluidos en el conjunto de Cuadros de Indicadores al final de la Sección I de este manual son de mediano plazo (por ejemplo, *No. de veces que un joven ha tenido una ITS en el último año*).

- *frecuencia de comunicación de los jóvenes con sus padres en la última semana.*
  - *tiempo promedio de la comunicación entre padres e hijos*
  - *temas tratados por los jóvenes y sus padres*
  - *percepción de los jóvenes sobre la calidad de la comunicación en la última semana*
2. Desarrollar indicadores que exploren las relaciones entre dos medidas del mismo fenómeno. Por ejemplo, al explorar la relación padre-hijo, se consideran dos indicadores relacionados:
- *La percepción de los jóvenes de si los padres los entienden*
  - *La percepción de los jóvenes de qué tipos de problemas pueden tratar con sus padres*

Midiendo estos dos indicadores se puede evaluar el grado en que se verifican relaciones hipotéticas entre conceptos relacionados. Por ejemplo, se puede medir si todos los jóvenes que dicen que sus padres los entienden también dicen que pueden hablarle a sus padres de una variedad de problemas.

Se puede aumentar la confiabilidad de los indicadores reduciendo la posibilidad de que ocurran condiciones de azar o temporales en una persona, situación o conjunto de procedimientos de medición.

1. Controlar la consistencia de las respuestas de un individuo haciéndole preguntas similares más de una vez durante una encuesta o entrevista. Por ejemplo, un joven que reporta una interacción de calidad con sus padres, pero también que no puede hablarle a sus padres cuando tiene problemas muestra inconsistencia en sus respuestas. En los análisis de datos se puede controlar cuántos jóvenes dieron respuestas inconsistentes. Si muchos jóvenes las dieron, se habría identificado una medida poco confiable de estos indicadores. Si sólo unos pocos la dieron, se habría identificado un error en la comprensión del individuo acerca de esas preguntas.
2. Recolectar datos en diferentes momentos y controlar cuán consistentes son las respuestas de los jóvenes. Por ejemplo, se puede hacer la misma serie de preguntas acerca de la calidad de la interacción con los padres en encuestas realizadas cada seis meses.
3. Evaluar los datos recolectados buscando inconsistencias debidas a errores en los procesos de observación, codificación y registro de datos. Por ejemplo, controlar si los jóvenes entrevistados por entrevistadores de distintas edades tienen respuestas significativamente diferentes. También controlar si las respuestas a preguntas abiertas están codificadas correctamente; por ejemplo, si “feliz” y “alegre” están codificados como la misma o distinta respuesta.

Probar rigurosamente la validez y confiabilidad puede requerir ayuda externa para efectuar pruebas estadísticas. Como mínimo, es importante considerar estos temas a medida que se desarrollan los indicadores.

| <p style="text-align: center;"><b>Hoja de trabajo 4.1</b><br/><b><i>Preparación de una lista de posibles indicadores</i></b></p>  |   |                           |                  |  |
|---|---|---------------------------|------------------|--|
| <p>1. Escribir los objetivos en el cuadro</p> <p>2. Para cada objetivo, escribir las actividades planificadas para lograr el objetivo. Referirse al Modelo Lógico desarrollado para asegurar que se incluyan actividades que aborden todos los factores antecedentes.</p> <p>3. Para cada actividad, anotar quién participará (por ejemplo, jóvenes entre 8 y 12 años; varones; poblaciones vulnerables) y dónde se llevará a efecto.</p> <p>4. Para cada actividad, referirse a los Cuadros de Indicadores para listar todos los posibles indicadores (Diseño de Programa, Desarrollo y Funcionamiento de Sistemas de Programas, Implementación de Programas, Resultados de la Intervención de Programas) o desarrollar indicadores propios.</p> |   |                           |                  |  |
| <b>Objetivos</b>  | <b>Actividades</b>                        | <b>Población objetivo</b> | <b>Ubicación</b> | <b>Posibles indicadores</b>                              |
| <i>Objetivo 1</i>   | Actividad 1<br>Actividad 2<br>Actividad 3 |                           |                  | Indicador 1<br>Indicador 2<br>Indicador 3<br>Indicador 4 |
| <i>Objetivo 2</i>   | Actividad 1<br>Actividad 2<br>Actividad 3 |                           |                  | Indicador 1<br>Indicador 2                               |
| <i>Objetivo 3</i>   | Actividad 1<br>Actividad 2                |                           |                  | Indicador 1<br>Indicador 2<br>Indicador 3                |
| <i>Objetivo 4</i>   | Actividad 1<br>Actividad 2                |                           |                  | Indicador 1<br>Indicador 2<br>Indicador 3                |

| Hoja de trabajo 4.2  |                      |   |  |   |  |   |                  |  |    |   |   |   |  |
|--|----------------------|---|--|---|--|---|------------------|--|----|---|---|---|--|
| <i>Evaluación de los posibles indicadores</i>  |                      |   |  |   |  |   |                  |  |    |   |   |   |  |
| 1. Listar los indicadores de la Hoja de Trabajo 4.1 en la primera columna.<br>2. Aclarar el alcance del programa. ¿Es un esfuerzo en gran escala para llegar a todos los miembros de la población objetivo o una intervención más pequeña, más limitada, que llegará sólo a aquellos que participan en servicios o actividades específicas?<br>3. Para cada indicador anote las posibles fuentes de datos necesarios, tales como encuestas, grupos focales o mapas.<br>4. Para cada fuente de datos, marcar si los datos están disponibles o necesitan ser recolectados.<br>5. Clasificar la dificultad de la recolección de datos basada en la disponibilidad, tiempo y costo de recolección.<br>6. Clasificar la importancia del indicador (alta o baja).<br>7. Determinar la prioridad basada en la dificultad de la recolección de datos y la importancia del indicador. |                      |   |  |   |  |   |                  |  |    |   |   |   |  |
| Posibles indicadores (de la Hoja de Trabajo 4.1)   | Alcance del programa |   | ¿Están disponibles los datos en la actualidad?<br>S= Sí<br>N= No |   | ¿Necesita recolectar nuevos datos?<br>S= Sí<br>N= No |   | Fuentes de datos | Facilidad de la recolección de datos<br>F= Fácil<br>FC= Factible<br>D= Difícil |    |   | Importancia del indicador<br>A= Alta<br>B= Baja |   | Prioridad (1 es más alta)<br>1=FA 2=FCA<br>3=DA 4=FB<br>5=FCB 6=DB |
| Indicador 1  | G                    | P | S  | N | S  | N |                  | F  | FC | D | A   | B |  |
| Indicador 2  | G                    | P | S  | N | S  | N |                  | F  | FC | D | A   | B |  |
| Indicador 3  | G                    | P | S  | N | S  | N |                  | F  | FC | D | A   | B |  |
| Indicador 4  | G                    | P | S  | N | S  | N |                  | F  | FC | D | A   | B |  |
| Indicador 5  | G                    | P | S  | N | S  | N |                  | F  | FC | D | A   | B |  |
| Indicador 6  | G                    | P | S  | N | S  | N |                  | F  | FC | D | A   | B |  |

## □ Capítulo 5: Diseños de evaluación para determinar los impactos del programa

Una mirada al capítulo:

Este capítulo se centra en la determinación del impacto del programa.

- *Ofrece una orientación y consideraciones sobre las necesidades de una evaluación de impacto*
- *Examina los diseños de estudios que se pueden usar para hacer una evaluación de impacto*
- *Considera los requisitos técnicos y los recursos disponibles antes de comenzar*
- *Presenta opciones para iniciar las evaluaciones cuando el programa ya está funcionando*

### ...¿Por qué se debe hacer una evaluación de impacto?

Este capítulo se centra en los tipos de evaluaciones de impacto que se pueden hacer. Una evaluación de impacto revelará la medida en que los cambios observados en los indicadores de resultados se deben a las actividades del programa. Si sólo se miden los cambios en los indicadores de resultados en la evaluación, los hallazgos pueden no ser totalmente creíbles porque:

- **Otros eventos o condiciones pueden contribuir a los cambios en los indicadores de resultados.** El programa es sólo uno entre muchos factores que pueden estar afectando los resultados. Por ejemplo, los cambios en las condiciones económicas u otros cambios sociales pueden influir sobre lo que piensan los jóvenes acerca de la edad aceptable para comenzar a tener relaciones sexuales o sobre el uso de preservativos. Otros programas pueden estar dirigidos al mismo público objetivo y estar realizándose al mismo tiempo. Estos tipos de eventos externos pueden hacer que los efectos del programa parezcan mayores o menores de lo que realmente son. Estos factores se llaman *eventos seculares*.
- **Cambios en los individuos que están siendo estudiados a lo largo del tiempo.** El crecimiento y desarrollo de los niños, o *maduración*, afecta sus actitudes y su físico, amenazando la validez interna de una evaluación que pretende vincular los cambios en los resultados a elementos como el conocimiento, actitudes o habilidades adquiridas en los programas de educación o promoción de la salud.
- **Los participantes del programa pueden estar predispuestos a resultados específicos.** Es posible que el programa atraiga a jóvenes que están predispuestos a los resultados positivos fomentados por las actividades del programa. Por ejemplo, el programa puede atraer más que nada a jóvenes con aspiraciones educacionales altas, que puedan estar menos inclinados a tener comportamientos arriesgados. Si ese es el caso, la simple medición de los cambios en los indicadores de resultados exageraría la efectividad del programa, ya que muchos de los participantes del programa habrían tenido resultados positivos aun sin estar expuestos al programa. Este problema se llama *sesgo de selección*.

### *...¿Cuáles son los tipos de diseños de estudio para evaluaciones de impacto?*

Los tres tipos más importantes de diseños de estudio para evaluaciones de impacto son:

- *experimentos al azar,*
- *cuasiexperimentos, y*
- *diseños no experimentales*

Estos diferentes diseños de estudio se presentan en más detalle en las siguientes secciones de este capítulo. En general, algunos factores que diferencian a un diseño del otro son:

- **Si se usa un grupo de “control” o “comparación”.** Un grupo de control o comparación es un grupo de personas, instalaciones o comunidades similares a las que reciben la intervención, pero que no han sido expuestas a la intervención. El propósito de un grupo de control o comparación es proporcionar un estimado de lo que habría pasado si no se hubiera implementado el programa. Un *grupo de control* se asigna aleatoriamente, mientras un *grupo de comparación* tiene características similares pero no se asigna aleatoriamente.
- **Cómo se asignan los participantes a los grupos de intervención y comparación.** En algunos estudios de evaluación, los participantes son asignados a grupos de intervención y grupos de control a través de asignaciones aleatorias. En otros, los grupos de control son seleccionados—no asignados aleatoriamente—para

igualar las características del grupo de intervención, con la excepción de su exposición a la intervención que está siendo evaluada.

- **El momento en que se recolectan datos en relación a la implementación del programa.** Una evaluación puede recolectar datos antes, durante y/o después de la implementación del programa.
- **La complejidad del análisis estadístico requerido.** Algunos diseños de estudio requieren un análisis estadístico altamente sofisticado.

A medida que se consideran los distintos diseños de estudio, hay que tener en mente los siguientes factores al seleccionar el más apropiado para medir el impacto del programa:

- **Asuntos éticos:** El estudio sólo debe efectuarse si éste se puede llevar a cabo éticamente. Si el estudio compromete los derechos individuales o niega al grupo de control la oportunidad de recibir un programa o servicios que hubieran recibido si no participaran en el estudio, se deberá cambiar o abandonar el diseño de estudio.
- **La importancia de poder demostrar impacto:** Si el apoyo de la comunidad y/o el financiamiento depende de demostrar un impacto, es mejor usar diseños más rigurosos.
- **Validez:** La validez se refiere a la capacidad de un diseño de estudio de medir el “verdadero” impacto de un programa o intervención. Los diseños de estudio más rigurosos son aquellos que son menos vulnerables a los elementos que atentan contra la validez. Dos de estos elementos—eventos externos y sesgos de selección—se mencionaron al comienzo de este capítulo. El recuadro titulado “Resumen de Diseños de Estudio para Evaluaciones de Impacto” resume las otras grandes amenazas a la validez en las evaluaciones de impacto de un programa, y proporciona una guía para seleccionar diseños de estudio que sean menos vulnerables a las amenazas a la validez, dadas las circunstancias específicas y los recursos disponibles.
- **Consideraciones acerca de los recursos:** La disponibilidad de fondos, personal, tiempo y equipos también influirá en la selección del diseño de estudio. La selección debe estar basada en una evaluación realista de los recursos disponibles; lo que en algunos casos es deseable con recursos óptimos, puede no ser factible en la realidad.

### *...Experimentos al azar*

Los experimentos al azar son el diseño de evaluación más confiable y válido. En evaluaciones que usan este diseño de estudio, los participantes son asignados aleatoriamente a un grupo que recibirá una intervención (llamado el grupo de intervención) o a un grupo de control que no recibirá la intervención. Esto se conoce como *asignación aleatoria*, y es la mejor manera de asegurar que los miembros de los grupos de intervención y control sean los mismos en términos de características vitales, tales como edad, educación, antecedentes familiares, actitudes y cualquier otro factor que pueda influir sobre los resultados buscados por el programa.

Más adelante se explican los dos tipos principales de experimentos al azar—diseños de grupos de control con pretest-posttest y diseños de grupos de control solamente con posttest.

## **El diseño más sólido es el de grupo de control con pretest-postest**

Con los diseños de grupos de control con pretest-postest, se podrá:

- **asignar aleatoriamente** a personas, instalaciones o comunidades a grupos experimentales,
- hacer mediciones **antes y después** de la intervención y
- medir el impacto definido como la diferencia entre el cambio en los indicadores de resultados para el grupo de intervención y el grupo de control

Este es el diseño más riguroso de todos, considerado el “estándar de oro” entre los especialistas en evaluaciones. Sin embargo, presenta algunos desafíos:

- *No siempre es factible asignar aleatoriamente a los sujetos a cada grupo experimental*
- *Para que los experimentos al azar sean efectivos, se necesita mantener las condiciones experimentales durante el transcurso del programa y mantener a un mínimo el número de otros programas introducidos durante el estudio a los grupos de intervención o de control*

## **A veces surgen preocupaciones éticas**

¿Es ético privar a algunas personas de los beneficios posibles de una intervención para evaluar un programa? Cuando los recursos son suficientes para atender a una población entera, esta decisión puede ser en realidad muy difícil de tomar. Sin embargo, en la mayoría de los lugares los recursos del programa no son suficientes para atender poblaciones enteras. En estas situaciones, se usan diseños de intervenciones “por fases” o “retardadas” para poder dar a todos los participantes la oportunidad de recibir intervenciones o servicios. Los diseños de programas por fases consisten en proporcionarle, en una etapa subsiguiente, la intervención al grupo de jóvenes que fueron estudiados como grupo de control en la primera fase del programa. Este diseño representa un avance importante en permitir que se realicen evaluaciones sólidas de una forma ética.

A continuación, se presentan dos ejemplos de estudios grandes, quizás más grandes de lo que es factible para algunos programas. Sin embargo, estos ejemplos muestran la factibilidad de evaluar programas de salud reproductiva.



## Evaluación de un programa para posponer la iniciación sexual: Un experimento al azar con pretest-postest

### ¿Qué tipo de programa se estaba evaluando?

Entre 1992 -1994 los evaluadores estudiaron el programa para Posponer la Iniciación Sexual (PIS) en California, EE.UU., usando un diseño al azar de grupo de control con pretest-postest.

### ¿Qué temas se estaban midiendo con los indicadores de resultados?

- creencias acerca de la actividad sexual,
- razones para abstenerse o tener relaciones sexuales,
- creencias acerca del sexo y los medios de comunicación,
- comunicación personal,
- autoeficacia (por ejemplo, confianza en la capacidad de uno mismo para negociar el uso de preservativos cuando sea necesario),
- intenciones del comportamiento y
- comportamientos reales

### ¿Cómo se asignaron los grupos experimentales?

Un total de 10.600 jóvenes fueron asignados a grupos experimentales de la siguiente manera:

- En algunos sitios, se asignaron al azar salas de clases dentro de las escuelas de la muestra a grupos de intervención dirigidos por jóvenes, dirigidos por adultos o grupos sin intervención (que sirvieron de control).
- En otros sitios, se asignaron al azar escuelas completas a grupos de intervención (guiados por adultos) o a grupos sin intervención (de control).
- Se reclutaron jóvenes de organizaciones basadas en la comunidad que fueron asignados al azar, ya sea a grupos de intervención (guiados por adultos) o a grupos sin intervención (de control).

### ¿Cuándo se recolectaron los datos?

Los datos de la encuesta fueron recolectados en tres momentos: antes de la implementación del programa, 3 meses después de la implementación y 17 meses después de la implementación.

### ¿Cómo se midió el impacto?

El impacto se midió comparando los cambios en los indicadores de resultados para los jóvenes en cada uno de los tres conjuntos de grupos experimentales, a los 3 meses y nuevamente a los 17 meses después de la implementación del programa.

### ¿Cuál fue el impacto?

El programa no influyó en el conocimiento, las creencias, ni el comportamiento que se intentó cambiar a través de la intervención. Surgieron algunas tendencias en los indicadores a los tres meses, pero éstas no se manifestaron a los 17 meses.

Adaptado de Kirby y cols., 1997.

## Prevención del VIH en Tanzania: Un experimento al azar con pretest-postest

### ¿Qué tipo de programa se estaba evaluando?

Entre 1991-1994 los evaluadores estudiaron un programa de reducción del VIH en Mwanza, Tanzania, cuya meta era mejorar la detección y tratamiento de las infecciones de transmisión sexual (ITS). Esta evaluación, que incluyó un total de 12.537 personas, usó un diseño aleatorio de grupo de control con pretest-postest.

### ¿Qué áreas se estaban midiendo con los indicadores de resultados?

- tasas de conversión serológica para el VIH,
- prevalencia de ITS y
- prevalencia de comportamientos sexuales de riesgo seleccionados.

### ¿Cómo se asignaron los grupos experimentales?

La asignación al azar de la población del estudio se hizo de la siguiente manera:

- Para el estudio, se eligieron seis pares de comunidades (12 comunidades en total) semejantes (en términos de ubicación y factores socioeconómicos).
- En cada uno de estos seis pares se asignó al azar una comunidad al grupo de intervención y otra al grupo de control.
- En cada comunidad, se eligió una muestra al azar de 1.000 hogares, y todas las personas de 15 años o más fueron incluidas en la muestra.

### ¿Cuándo se recolectaron los datos?

Se recolectaron datos biológicos y de encuestas en dos momentos: antes de la implementación del programa y nuevamente 24 meses después.

### ¿Cómo se midió el impacto?

El impacto se midió comparando los cambios en los indicadores de resultados para las personas que viven en las aldeas "intervenidas" comparadas con aquellas en las aldeas de control, durante el período de los 24 meses del estudio.

### ¿Cuál fue el impacto?

Las tasas de infección del VIH en los 24 meses del estudio resultaron más bajas en las aldeas que recibieron servicios de detección y tratamiento de ITS, comparadas con las aldeas donde no hubo tales servicios.

Adaptado de Grosskurth y cols., 1995.

## El diseño de grupo de control solamente con posttest es algo débil

Con el diseño de grupo de control solamente con posttest se podrá:

- **asignar al azar** a personas, instalaciones o comunidades a grupos experimentales (igual que el diseño de grupo de control con pretest-posttest),
- **medir indicadores sólo después** de la intervención, y
- **medir el impacto como la diferencia entre los indicadores de resultados** para el grupo de intervención y el grupo de control, en algún momento después de la implementación del programa

El supuesto clave de este diseño es que los grupos experimentales son equivalentes en términos de todos los factores excepto la exposición a la intervención.

El diseño de grupos de control solamente con posttest es más débil que el diseño de grupos de control con pretest-posttest porque el primero no hace mediciones antes de la intervención de un programa. Los diseños de grupos de control solamente con posttest no permiten conocer la situación de cada grupo antes de las intervenciones de los programas. Si bien se pueden mostrar algunas diferencias entre los participantes del programa y el grupo de control después de la implementación, tener medidas de su situación antes de la intervención permite obtener conclusiones más sólidas sobre los efectos del programa.

### Lista de verificación

¿Podemos usar un diseño de estudio con asignación al azar si...?

- La intervención que queremos evaluar no ha empezado aún.
- El personal u otros profesionales tienen la capacidad de efectuar asignaciones al azar.
- Existe un grupo de personas que pueden ser asignadas al azar para recibir la intervención y otro grupo para participar en el estudio como grupo de control.
- Los recursos financieros y la capacidad técnica de recolectar datos antes y después que comience la intervención están disponibles.
- No existen problemas éticos en negar el programa a los grupos de control durante el estudio.

## *...Cuasiexperimentos*

### Los cuasiexperimentos usan grupos experimentales similares, seleccionados a través de métodos que no son al azar

Aunque no sea factible asignar grupos experimentales al azar, se pueden controlar muchos de los factores externos que afectan a los grupos de control e intervención usando un diseño *cuasiexperimental*. Esto se puede hacer eligiendo un grupo de control que sea lo más parecido posible al grupo de intervención, a menudo haciendo coincidir las características que se consideran antecedentes importantes de los resultados que busca un programa.

Por ejemplo, si se está ejecutando y evaluando un programa basado en las escuelas, se podrían elegir escuelas de control que sean similares a las escuelas de intervención, pero que no son parte del programa. Algunas de las características que se pueden considerar al elegir las escuelas de control son:

- *ubicación geográfica,*
- *número de grados que se enseñan en cada escuela,*
- *composición socioeconómica de la población estudiantil y*
- *nivel de capacitación y experiencia de los profesores.*

Las escuelas de control deberían entonces coincidir con las escuelas de intervención en cada una de estas características.

### **Compensar por las fallas en el proceso de hacer coincidir las características**

Debido a que hacer coincidir las características generalmente se ve como un sustituto imperfecto de la asignación al azar, generalmente en los cuasiexperimentos se usa el análisis estadístico para tomar en cuenta o “controlar” las diferencias en los factores que no se pudieron tomar en cuenta a través de la coincidencia de las características. Generalmente, al hacer coincidir las características sólo se pueden tomar en cuenta algunos de los muchos factores en los que pueden diferir los grupos de comparación.

Las siguientes secciones explicarán los tres tipos de cuasiexperimentos usados más comúnmente:

- *diseño de grupos de control no equivalentes con pretest-postest,*
- *diseño de grupos de control no equivalentes solamente con postest y*
- *diseño de control genérico.*

### **El diseño de los grupos de control no equivalentes con pretest-postest es un tipo de cuasiexperimento usado ampliamente**

Este es quizás el diseño más usado en evaluaciones de impacto de programas sociales. Es similar al diseño al azar de grupos de control pretest-postest, excepto que el grupo de comparación se crea haciendo coincidir características en vez de hacerlo por asignación al azar, por lo que se llama “no equivalente”. Los miembros del grupo no se asignan al azar y la probabilidad de ser asignados al grupo de intervención o al de control no es la misma.

Con el diseño de grupos de control no equivalentes con pretest-postest se podrá:

- *crear grupos experimentales **haciendo coincidir** características específicas que se consideran antecedentes importantes de los resultados que busca el programa,*
- ***medir indicadores antes y después** de la intervención y*
- *medir el impacto como la **diferencia entre los cambios en los indicadores de resultados** para el grupo de intervención y el grupo de control.*

El mayor riesgo de este diseño es que los grupos experimentales pueden diferir unos de otros en factores que no se pueden contabilizar a través de la coincidencia de características o un análisis estadístico. Por ejemplo, los grupos de comparación pueden diferir en factores que son difíciles de medir (por ejemplo, qué proporción había experimentado abuso sexual) o que simplemente no se midieron.

## Evaluación de un servicio de salud vinculado a las escuelas en Brasil

### *Un cuasiexperimento con grupo de control no equivalente con pretest-postest*

#### **¿Qué tipo de programa se estaba evaluando?**

En 1995, en la ciudad de Salvador (Bahía), en Brasil, los Secretariados de Salud y Educación comenzaron pruebas piloto para un programa de servicio de salud vinculado a las escuelas, las cuales intentaban reducir las tasas de embarazo y de ITS entre los jóvenes. El programa destacaba la educación sexual en las escuelas, las mejoras en los servicios en las clínicas de salud pública para hacerlos más “afines a los jóvenes” y los vínculos formales entre las escuelas y los servicios de salud.

#### **¿Qué temas fueron medidos por los indicadores de resultados?**

- conocimientos y actitudes
- comportamiento sexual
- desarrollo integral de los jóvenes
- factores de riesgo ambientales y protectores

#### **¿Cómo se asignaron los grupos experimentales?**

Se usó un diseño cuasiexperimental con pretest-postest para medir el impacto del programa. Se eligieron al azar seis escuelas en el área metropolitana de Salvador para el grupo de intervención y seis escuelas de control de iguales características. Se hizo coincidir a las escuelas en términos de ubicación, número de grados de estudio y nivel socioeconómico de los alumnos.

#### **¿Cuándo y de dónde fueron recolectados los datos?**

- Los datos fueron recolectados en escuelas intervenidas y de control antes de la implementación del programa y tres años después.
- Se efectuó una encuesta de “monitoreo” en las escuelas intervenidas al final del primer año del programa. Los estudiantes que se cambiaron de escuela durante este período fueron seguidos a sus nuevas escuelas.
- Se realizaron encuestas en los servicios de salud—incluyendo observaciones de transacciones de servicios y entrevistas de salida con los clientes jóvenes—a la mitad de la implementación del programa para evaluar la “amigabilidad” de los servicios a los jóvenes.

#### **¿Cuál fue el impacto?**

Se medirá el impacto a través de indicadores de actitud responsable en comportamiento sexual, habilidad de negociar relaciones y uso de servicios de salud reproductiva.

Fuente: Gaffikan, Lynne, Estela Maria Leao de Aquino, Robert Magnani, Maria de Conceicao Chagas Almeida, Eric Seiber, Amara Robinson y Varja Lipovsek, "Evaluation of an Integrated Adolescent Sexuality Education/Health Service Provider Training Pilot Project in Salvador, Bahia, Brazil, Focus On Young Adults." *FOCUS on Young Adults*, Washington, D.C., con las Secretarías de Salud y Educación del Estado de Bahía, el Instituto de Salud Comunitaria de la Universidad Federal de Bahía y la corporación JHPIEGO. Octubre de 2000.

**El diseño de grupo de control no equivalente solamente con posttest (comparación de grupo estático) es un tipo de cuasiexperimento algo débil**

Este diseño es similar al de grupo de control no equivalente con pretest-posttest, excepto que no se incluye un pretest.

Con el diseño de un grupo de control no equivalente solamente con posttest se puede:

- *crear grupos experimentales pareando características específicas que se consideran antecedentes importantes de los resultados que busca el programa,*
- **medir indicadores solamente después** de la intervención y
- *medir el impacto como la **diferencia entre los indicadores de resultados** para los grupos de intervención y de control en algún momento posterior a la intervención del programa.*

Este diseño se usa a menudo cuando no se recolectó información inicial. Sin embargo, es un diseño relativamente débil, ya que existe un alto riesgo de que los grupos experimentales puedan diferir en aspectos que no se pueden contabilizar a través del pareo o el análisis estadístico. Cuando se usa este diseño, se necesitan técnicas estadísticas más sofisticadas para producir estimados defendibles del impacto del programa. Sin embargo, incluso cuando se usen técnicas estadísticas más sofisticadas, existe el peligro de que subsistan las diferencias entre los grupos experimentales, las cuales no fueron medidas durante la recolección de datos. Por esta razón, este diseño a menudo produce resultados que son vistos con escepticismo.

**Evaluación de un programa de educación sexual en escuelas del Perú:  
*Un grupo de control no equivalente solamente con posttest***

**¿Qué tipo de programa se estaba evaluando?**

En 1995, el Ministerio de Educación del Perú inició un programa nacional de educación sexual en las escuelas secundarias. Debido a la limitación de recursos, el programa se efectuaría por etapas durante un período de 5 a 10 años. Debido a una variedad de razones, no fue posible efectuar una encuesta base antes de la implementación del programa.

En 1998, el Ministerio decidió evaluar el programa. Debido a que no se había efectuado ninguna encuesta base, se eligió un diseño de grupo de control solamente con posttest para la evaluación.

**¿Cómo se asignaron los grupos experimentales?**

Se eligieron al azar 22 escuelas que habían implementado el programa en Lima y otras ciudades seleccionadas en el interior del país. Se eligió una escuela de "control" con iguales características para cada par de escuelas "intervenidas", resultando en una muestra de 33 escuelas. Los criterios usados para hacer coincidir a las escuelas intervenidas con las escuelas de control incluyeron ubicación geográfica, tamaño, antigüedad del colegio y nivel socioeconómico del cuerpo estudiantil.

**¿Qué medían los indicadores de resultados?**

- conocimientos y actitudes
- comportamientos arriesgados y conductas de búsqueda de salud

### **¿Qué métodos de recolección de datos se usaron?**

Se recolectó información de estudiantes, profesores y administradores en cada escuela de la muestra, utilizando cuestionarios autoadministrados. También se efectuaron entrevistas con grupos focales de estudiantes y profesores.

### **¿Cuál fue el impacto?**

No se observaron diferencias consistentes entre los grupos con respecto a conocimientos, creencias o comportamientos.

Fuente: *Pathfinder/Perú*, con el Ministerio de Educación de Perú, "Evaluation of Results, 1996-1998 Program Period and Baseline Survey, 1999-2002 Program Period of the National Program of Family and Sexual Education, Ministry of Education (Peru)". Octubre de 1999. De *Pathfinder/Perú*: Delicia Ferrando, Dorina Vereau, Luisa Kanashiro y Carlos Cáceres; del Ministerio de Educación de Perú: Carmen Amelia Ríos de Coloma, José Vallejos, Víctor Montero y César Londona.

### **El tercer tipo de cuasiexperimento, los diseños de control genérico, comparan cambios en los indicadores de resultados con la situación de la población general**

Los diseños de control genérico se pueden utilizar para evaluar si los cambios o tendencias en los indicadores de resultados para los jóvenes expuestos al programa son diferentes a los de la población general de jóvenes. Por ejemplo, se pueden comparar datos de un programa con datos de una encuesta nacional acerca del uso de preservativos entre jóvenes. Utilizando un diseño de control genérico se podría establecer que la proporción de la población objetivo del programa que usó preservativos en su último encuentro sexual aumentó significativamente durante el transcurso del programa, pero que aún es menor que la proporción registrada entre los jóvenes a nivel nacional.

### **Encontrar datos relevantes para la población general**

Se puede usar diseños de control genérico sólo si existen datos respecto a las tendencias en los indicadores de resultados *relevantes* para la población general de jóvenes. En muchos países es posible encontrar este tipo de datos en encuestas de gran escala (por ejemplo, Encuestas Demográficas y de Salud, Encuestas de Salud Reproductiva de los Adolescentes) que se hacen periódicamente (por ejemplo, más o menos cada cinco años). Estas encuestas generalmente miden una serie de indicadores de resultados relevantes para los programas de salud reproductiva de los adolescentes, tales como edad al momento de la iniciación sexual, frecuencia de las relaciones sexuales, número de parejas sexuales, conocimientos acerca de factores de riesgo para el VIH/SIDA, uso de preservativos, etc.

## ¿Cuándo se debe usar un diseño de estudio cuasiexperimental?

Si no existen las condiciones para usar un diseño de estudio con asignación aleatoria.

### **Para todos los tipos de cuasiexperimentos:**

- Se han identificado las características de los jóvenes que son importantes para hacer coincidir a los grupos de intervención y de comparación.
- Existe un grupo de jóvenes que tienen las mismas características que el grupo de intervención que interesa evaluar.
- El personal u otros profesionales están capacitados para evaluar la coincidencia de las características.
- No surgen problemas éticos al negar el programa al grupo de comparación durante el estudio.

### **Para el diseño con pretest-postest:**

- Aún no ha comenzado la intervención.
- Se dispone de los recursos financieros y habilidad técnica para recolectar datos antes y después de que comience la intervención.

### **Para el diseño solamente con postest:**

- La intervención ha comenzado o la recolección inicial de datos no es factible por alguna otra razón (tal como falta de recursos).
- El personal u otros profesionales pueden manejar análisis estadísticos sofisticados.

### **Para diseño de control genérico:**

- Existen datos acerca de jóvenes que tienen características similares a aquellos que participan en el grupo de intervención bajo estudio.
- Los datos disponibles miden el tipo de indicador de resultado que la intervención está tratando de influir.
- Los datos del grupo de control genérico serán recogidos más o menos al mismo tiempo que se recolectarán los datos del grupo de intervención.



## **Condiciones para el uso del diseño de control genérico**

Si se quiere usar el diseño de control genérico para medir el impacto del programa, hay que asegurarse que:

- *la población para la cual existen datos de control disponibles es similar en características y composición a la población objetivo del programa. Este es el desafío más importante en el uso de los diseños de control genérico. Por ejemplo, si el programa tiene como objetivo a la población de jóvenes de alto riesgo, puede que no sea apropiado usar a la población general como grupo de comparación,*
- *los indicadores de resultados obtenidos en la población de control genérico son relevantes para el programa y*
- *los datos de encuestas para ambos grupos pueden obtenerse más o menos al mismo tiempo.*

## ***...Diseños no experimentales***

### **Los diseños no experimentales no usan grupos de control o de comparación**

Los diseños no experimentales no usan grupos de control, por lo tanto es una forma generalmente más débil de medir el impacto del programa que los diseños experimentales. Los diseños no experimentales se usan cuando:

- *no se ha previsto el uso de grupos de control o de comparación como parte del plan de evaluación o*
- *se espera que un programa o intervención “llegue” a toda la población objetivo juvenil; este tipo de programa a menudo se llama programa de cobertura total (por ejemplo, campañas intensivas de medios de comunicación masivos y programas nacionales de educación sexual en lugares donde existe una alta proporción de jóvenes que permanece en la escuela hasta el nivel secundario)*

Los siguientes párrafos describen tres tipos de diseños no experimentales:

- *diseño con mediciones periódicas,*
- *diseño con pretest-postest y*
- *diseño con postest solamente.<sup>32</sup>*

---

<sup>32</sup> Estos son clasificados por algunos como diseños cuasi experimentales, pero debido a que no tienen grupos de comparación, se clasifican aquí como diseños no experimentales.

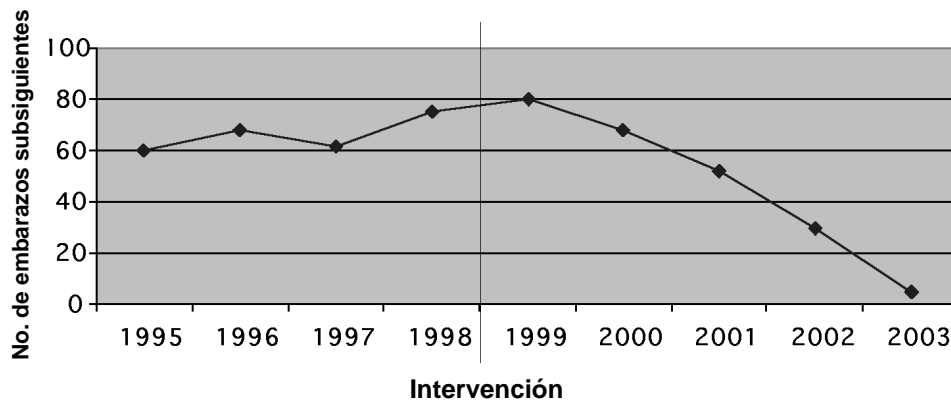
## Los diseños con mediciones periódicas usan las tendencias preintervención como comparación

El diseño seriado es el más riguroso de los diseños no experimentales. Con este diseño se podrá:

- *tomar varias medidas de indicadores de resultados para la población objetivo de un programa antes y después de la intervención,*
- *usar indicadores generados con anterioridad a la implementación para proyectar que habría pasado en la ausencia de un programa (asumiendo que esta tendencia hubiera continuado en ausencia de un programa),*
- *comparar la tendencia de los indicadores antes de la intervención a las tendencias después de la intervención y*
- *detectar el impacto del programa como un cambio de tendencias después de la implementación de una intervención (ver recuadro que sigue).*

### Análisis ilustrativo con mediciones periódicas

En este ejemplo, se midió el número de embarazos subsiguientes en la población objetivo de jóvenes entre 15 y 19 años en un área geográfica específica durante un período de cuatro años antes de que comenzara la intervención en 1999. Los datos recogidos mostraron que un año después de comenzada la intervención, el número de embarazos subsiguientes comenzó a disminuir. Esta tendencia a la baja se puede asociar a la intervención



## El diseño con mediciones periódicas tiene desventajas

Existen dos desventajas importantes en el diseño seriado:

- *Para usar este diseño, se deben completar un número suficiente de observaciones antes de implementar el programa (por lo menos tres, y preferiblemente más), para poder establecer la tendencia que habría ocurrido ante la ausencia de un programa. (Debido a que este tipo de datos preintervención rara vez está disponible, el diseño con mediciones periódicas no se usa ampliamente).*

- *Es difícil contabilizar los efectos de la maduración adolescente y otros factores externos que pueden haber influido en los resultados durante el período en que se implementó el programa.*

### **El diseño con pretest-postest es un tipo más común de diseño no experimental**

El diseño con pretest-postest se usa comúnmente cuando no se tiene un grupo de control, y no se tienen suficientes datos preintervención para usar un diseño con mediciones periódicas. En este diseño se puede:

- *medir los indicadores de resultados para la población objetivo de un programa, tanto antes como después de la intervención y*
- *probar hipótesis sobre los cambios observados para ver su significación estadística (es decir, para determinar si los cambios fueron mayores de lo que se podía esperar a través de errores aleatorios de medición)*

### **Contratación de consultores externos para ayudar con las evaluaciones**

Contratar consultores externos para ayudar con el esfuerzo de evaluación puede ser una buena manera de disminuir las exigencias al personal, darle credibilidad al esfuerzo de evaluación y compensar por las habilidades que el personal del programa no tiene. Para asegurarse de que el consultor le proporcione al programa la ayuda que realmente necesita, deben seguirse los siguientes consejos:

- Determinar, con el personal y los interesados, las tareas que efectuará el consultor y especificar claramente el alcance del trabajo. Listar las habilidades y fortalezas que debería tener el consultor a ser contratado.
- Entrevistar a varios candidatos, y asegurarse de que la persona contratada tiene las habilidades necesarias y una visión compartida de lo que logrará la evaluación.
- Hacer un contrato que deje claras las expectativas, especificando los derechos del consultor en el uso de los datos recolectados para la evaluación.
- Antes de que el consultor comience a trabajar, presentarlo al personal y los interesados para debatir los temas de su interés y sus preocupaciones.
- Designar a un miembro del programa para que actúe de vínculo entre el programa y el consultor.
- Si el consultor efectuará toda la evaluación, asegurarse de que rinda cuentas de los hallazgos de la evaluación al personal y a los interesados antes de publicar los resultados.

### **Dificultades para atribuir el impacto al programa**

Aunque este diseño permite documentar los cambios en los indicadores de resultados en la población objetivo del programa, es difícil saber hasta qué punto estos cambios se deben en realidad al programa. Factores externos podrían tener efectos positivos o negativos en los resultados esperados del programa, escondiendo así los verdaderos efectos del programa. Si hay tendencias positivas en los indicadores de resultados, lo más que se puede decir es que el programa puede haber contribuido a los cambios positivos. En el caso de hallazgos negativos en la evaluación de impacto, apenas se podrá negar la conclusión de que el programa fue inefectivo—aun cuando el impacto negativo pudo haber sido causado por factores externos.

## **Uso de análisis sofisticado para fortalecer el diseño con pretest-postest**

Una manera de fortalecer los hallazgos de los diseños con pretest-postest es buscar *asociaciones estadísticas* entre el nivel de exposición de los sujetos del estudio al programa y los cambios en los indicadores de resultados. Por ejemplo, si se puede demostrar que los jóvenes con altos niveles de exposición al programa tuvieron resultados más favorables que aquellos con menos exposición o sin exposición, se puede presentar un caso más sólido en favor del impacto del programa. Lo mismo ocurre si se puede demostrar que hubo indicadores de resultados más favorables en los lugares del programa que exhibieron mejores indicadores de implementación del programa que en los lugares con peores indicadores de implementación.

Estos tipos de análisis pueden ser complicados. Si se usa un diseño no experimental de este tipo, se debería buscar ayuda de un consultor externo, firma de investigación o profesor universitario para efectuar los análisis.

### **Educación de pares en Paraguay: *Un diseño no experimental con pretest-postest***

#### **¿Qué tipo de programa se estaba evaluando?**

Entre 1997-1999 se implementó en Asunción y en la región central de Paraguay, un programa de educación de pares basado en los medios de comunicación. El objetivo del programa era reducir las tasas de embarazo y de ITS entre los jóvenes. En el programa destacaban una presentación radial, actuaciones artísticas y diseminación de mensajes relevantes a través de otros canales de comunicación. Los adolescentes educadores eran los actores/presentadores en todas las actividades.

Debido a que el programa era un programa de cobertura total (es decir, se esperaba “llegar” a todos los jóvenes en Asunción y la región central), no fue posible hacer un diseño de estudio experimental. Por lo tanto, se usó un diseño no experimental.

#### **¿Cuándo se recolectaron datos y a través de qué métodos?**

Los cambios en los indicadores de resultados se midieron a través de encuestas efectuadas entre jóvenes en la escuela y fuera de ella, antes de la implementación del programa y dos años después de su implementación.

#### **¿Qué se estaba midiendo con los indicadores de resultados?**

- conocimientos y conciencia de los temas de salud reproductiva entre los jóvenes
- cómo los medios de comunicación tratan los temas de SSRA

#### **¿Cómo se analizaron los datos?**

Para evaluar si los cambios observados en los resultados podrían ser atribuidos al programa, se relacionaron los indicadores de comportamiento obtenidos de la encuesta de seguimiento con el nivel de exposición de los sujetos a las actividades del programa, eventos y mensajes. Se usaron métodos estadísticos sofisticados para controlar si hubo otras diferencias relevantes entre los grupos con alto y bajo nivel de exposición al programa.

#### **¿Cuál fue el impacto?**

Se demostraron diferencias significativas en conocimiento, creencias y comportamientos en el grupo más expuesto que en el menos expuesto al programa. Estas diferencias se atribuyeron al programa. Por ej., un incremento en la percepción de que las jóvenes que se “protegen” sexualmente son más responsables. También se observó un aumento en la proporción de jóvenes que usaron preservativos en su primer coito, pero este aumento no se pudo atribuir al programa.

Adaptado de *FOCUS on Young Adults*, informe final a ser publicado.

## **El diseño con posttest solamente es el más débil de los diseños no experimentales**

En el diseño con posttest se recogen los datos para una población objetivo sólo después de que se ha terminado la intervención. Este es el más débil de los métodos para medir el impacto del programa, ya que no permite medir:

- *cambios en los indicadores de resultados a lo largo del tiempo, ni*
- *las diferencias con un grupo de jóvenes no expuestos al programa*

Debido a las debilidades mencionadas, muchos especialistas en evaluación ni siquiera llamarían a este un diseño de estudio. Sin embargo, este diseño puede ser útil para medir aspectos aparte del impacto del programa, tales como la cobertura del programa, las características de los jóvenes que han o no han sido servidos por el programa, y las razones por las cuales los jóvenes no utilizan el programa.

### **¿Cuándo se debe utilizar un diseño de estudio no experimental?**

...cuando no existen las condiciones para usar una asignación aleatoria o un estudio de diseño cuasiexperimental.

#### **Para todo tipo de estudios no experimentales (sólo necesita marcarse una razón):**

- Es financiera o logísticamente imposible recolectar datos de un grupo de comparación.
- La intervención está llegando a toda la población objetivo de jóvenes.
- Éticamente la intervención no se puede negar a los grupos de comparación durante el estudio.

#### **Para el diseño con mediciones periódicas:**

- La intervención aún no ha comenzado.
- Se tienen los recursos financieros y la capacidad técnica para recolectar datos por lo menos tres veces antes de que comience la intervención.
- El personal y otros profesionales tienen la habilidad de proyectar tendencias basadas en los datos base.

#### **Para el diseño con pretest-posttest:**

- No es factible recolectar datos base del grupo de intervención.
- El personal y otros profesionales tienen las habilidades necesarias para efectuar análisis estadísticos más sofisticados.

## *...Estudios de panel*

### **Los estudios de panel usan el mismo conjunto de sujetos de estudio para cada ronda de recolección de datos**

Los estudios de panel pueden efectuarse como parte de cualquier otro diseño de estudio. Un *estudio de panel* recolecta datos de los mismos sujetos en cada ronda de recolección de datos. Por ejemplo, si se están midiendo cambios en los indicadores de resultados de comportamiento en un programa basado en la escuela, se pueden recoger datos del mismo “panel” de estudiantes en dos o tres oportunidades. A pesar de que los estudios de panel a menudo involucran más de dos rondas de recolección de datos, cualquiera de los diseños con pretest-postest descritos anteriormente pueden considerarse un estudio de panel si los datos se recogieron del mismo “panel” de sujetos de estudio en la encuesta base y las de seguimiento.

La alternativa a usar un estudio de panel es seleccionar una nueva muestra de la población estudiada en cada ronda de recolección de datos. Por ejemplo, en vez de recolectar siempre datos de los mismos estudiantes en una evaluación de un programa basada en la escuela, se podría elegir una nueva muestra de estudiantes dentro de la misma muestra de escuelas en cada ronda de recolección de datos.

### **Existen ventajas y desventajas en los diseños de panel**

Si bien los diseños de panel permiten medir los cambios en los indicadores con mayor precisión que en los diseños que no son de panel, también pueden ser difíciles y costosos. Seguir a un panel de sujetos de estudio—especialmente jóvenes de alto riesgo—a lo largo del tiempo, durante un período largo de observación (por ejemplo, dos a tres años o más) requiere paciencia y recursos. Si se pierden muchos de los participantes de la muestra original, puede haber sesgos en los resultados.

En general, sólo se debe usar un diseño de panel cuando:

- *se tiene buena probabilidad de retener a una alta proporción de los participantes durante el período de ejecución esperado del estudio y*
- *se dispone de suficientes recursos para darle seguimiento a los participantes en el estudio en futuras rondas de recolección de datos.*

#### **¿Se debe usar un diseño de estudio de panel?**

- Si se quiere medir los cambios en algunos indicadores con mayor precisión.
- Si existe la probabilidad de retener a una alta proporción de los participantes en el estudio durante el programa (por ejemplo, los participantes en el estudio no son altamente móviles o difíciles de ubicar).
- Si se tiene suficientes recursos para seguir a los participantes en el estudio en futuras rondas de recolección de datos.

## Consejos para seleccionar un diseño de estudio

Los siguientes puntos ofrecen pautas para seleccionar un diseño de estudio en orden descendiente de mayor a menor validez:

- Tratar de crear grupos de intervención y de control asignando aleatoriamente a los miembros de la población objetivo en uno u otro grupo (experimentos al azar).
- Si la asignación aleatoria no es posible, tratar de encontrar un grupo de comparación que sea lo más parecido posible al grupo de intervención. Ver si el uso de control genérico es posible (cuasiexperimento).
- Si no hay disponibilidad de un grupo de control asignado al azar ni un grupo similar de comparación, tratar de usar un diseño con mediciones periódicas que pueda proporcionar información acerca de las tendencias antes y después de la implementación del programa (diseños no experimentales).
- Si no hay suficientes datos disponibles antes de la implementación para hacer un diseño con mediciones periódicas, tratar de obtener información base (pretest) que pueda ser comparada con la información posprograma (diseño no equivalente con pretest-postest).
- Si no se puede recolectar información base, estar consciente de que la capacidad de obtener conclusiones sólidas acerca del impacto del programa será limitada. Casi con seguridad habrá que usar técnicas de análisis estadístico sofisticadas.

Adaptado de Fischer y cols., 1991.

## ...Minimizar las amenazas a la validez de la evaluación

Si bien los diseños de estudio más rigurosos generalmente son más válidos, existen amenazas potenciales a la validez de una evaluación, sin importar qué diseño de estudio se use. A continuación se presentan las amenazas más importantes a la validez de las evaluaciones de impacto y sugerencias sobre cómo minimizarlas.

### Cómo minimizar las amenazas a la validez

| Amenaza a la validez   | Cómo minimizar la amenaza   |
|--|---|
| <p><b>La historia o eventos externos</b> que ocurren durante el período de la intervención pueden influir sobre el resultado de la intervención. Los cambios en factores esenciales de la sociedad (tales como condiciones económicas, políticas o sociales), condiciones climáticas extremas, cambios en las políticas públicas o factores específicos al programa pueden afectar los resultados. Si por ejemplo, una adolescente educadora quedara embarazada y no puede continuar con sus funciones como modelo apropiado del programa, eso puede mandar el mensaje a la población objetivo de que las relaciones sexuales sin protección son aceptables, comprometiendo la capacidad del programa para producir los resultados deseados.</p> | <ul style="list-style-type: none"><li>• Seleccionar un diseño de estudio con un grupo de control o de comparación.</li><li>• Monitorear los eventos externos, tales como los efectos de los medios de comunicación masivos, que pueden influir en la población objetivo para explicar los hallazgos inesperados de la evaluación.</li><li>• Efectuar una evaluación de proceso para entender cómo está funcionando el programa y, durante el análisis de los datos, poder tomar en cuenta los factores específicos del programa que pueden haber afectado los resultados.</li></ul> |

| Amenaza a la validez  | Cómo minimizar la amenaza   |
|---|---|
| <p>El <b>sesgo de selección</b> resulta de una diferencia entre las personas elegidas para la intervención y aquellas elegidas para el grupo de control. Por ejemplo, los jóvenes que viven cerca de un centro de atención específico pueden ser más adinerados, mejor educados y más involucrados en las actividades de la escuela o de clubes privados, por lo tanto, pueden no participar activamente en las actividades de ese centro. En cambio, los jóvenes que viven cerca de otro centro que no asisten a la escuela pueden tener menos oportunidades de recreación, por lo tanto, usan ese centro más frecuentemente. Esta diferencia podría significar que los que están expuestos a la intervención serían de un nivel socioeconómico o educacional diferente a los del grupo de control y, por lo tanto, responderían en forma diferente a la intervención.</p> | <ul style="list-style-type: none"> <li>• Elegir de un listado de hogares una muestra al azar de participantes en el estudio.</li> <li>• Consultar otros conjuntos de datos acerca de los jóvenes y comparar los antecedentes sociodemográficos de los jóvenes que participan en la intervención con aquellos jóvenes que no participan, e informar sobre los hallazgos en el informe de evaluación.</li> <li>• Asegurarse de hacer coincidir el máximo de características posibles entre el grupo de control y el grupo de intervención.</li> </ul> |
| <p>1. <b>Las pruebas</b> antes de la intervención (ya sea pruebas piloto de un instrumento de encuestas o pretest base) probablemente afectarán las respuestas que se den en el postest. Las personas que reciben un pretest (por ejemplo, jóvenes entrevistados antes de una campaña en los medios de comunicación enfatizando el sexo seguro) pueden recordar las preguntas y respuestas que dieron cuando respondan en el postest (después que se ha efectuado la campaña de los medios de comunicación). Las diferencias en los indicadores de preintervención y postintervención pueden deberse a la familiaridad con las preguntas, o a que los jóvenes se hayan “aprendido” las respuestas deseables, en vez de ser el resultado de la intervención.</p>   | <ul style="list-style-type: none"> <li>• Usar un grupo de control o comparación para que sea más probable identificar los cambios debidos a la intervención.</li> <li>• Probar los instrumentos de encuestas piloto con jóvenes que no participarán en el estudio.</li> <li>• Cambiar el orden, pero no el formato o contenido, de los puntos incluidos en el postest.</li> </ul>   |
| <p>2. <b>La maduración</b> ocurre con el pasar del tiempo y puede producir resultados en los participantes que no estén relacionados con los efectos del programa. Por ejemplo, si los participantes se cansan o se aburren durante un largo período de capacitación, los resultados de las pruebas pueden reflejar fatiga y distracción en vez de reflejar la efectividad de la capacitación. También, los jóvenes que están en el programa, a medida que maduran, pueden experimentar cambios en las actitudes y comportamientos que no están relacionados con el programa.</p>   | <ul style="list-style-type: none"> <li>• Asegurarse de que los administradores del programa estén conscientes de la necesidad de minimizar los efectos de la maduración lo más posible.</li> <li>• Usar un grupo de control o comparación para dar cuenta de la maduración.</li> <li>• Mantener la diversidad y vivacidad del programa, a la vez que se intente mantenerlo lo más interesante posible.</li> </ul>   |



| Amenaza a la validez  | Cómo minimizar la amenaza   |
|---|---|
| <p><b>La deserción</b> puede afectar los resultados de la evaluación. Si un programa se implementa por un período prolongado de tiempo, los participantes pueden desertar, graduarse, mudarse de ese lugar o morir y, por lo tanto, perderse para los estudios de seguimiento. Si los que se pierden son sustancialmente diferentes en términos de características significativas a los que se quedan, los resultados de un estudio de postintervención pueden reflejar esas diferencias en vez del efecto de la implementación del programa.</p>   | <ul style="list-style-type: none"> <li>• Comparar las características de los que continúan en el programa con las de los que desertan; si los dos grupos no son significativamente diferentes, se puede demostrar que desertar no afecta los resultados del estudio.</li> <li>• Asegurarse de recolectar información que permita contactar a los jóvenes estudiados a lo largo del tiempo. Si es posible, recolectar datos de alguna institución con la cual los jóvenes estén involucrados (por ejemplo, una escuela) y preguntar a cada entrevistado 1) dónde creen que estarán cuando se efectúe la encuesta de seguimiento y 2) la información para contactar una o dos personas que siempre sabrán donde están.</li> <li>• Presupuestar suficientes fondos para seguir a los jóvenes que se han mudado o han desertado.</li> </ul> |
| <p><b>Modificaciones</b> en la instrumentación, tales como modificar el cuestionario de una encuesta entre el pretest y el postest, pueden tener un efecto independiente de las intervenciones del proyecto. El deterioro de los equipos, tales como balanzas, reglas o calibradores de grasa pueden dar como resultado lecturas falsas. Algunos cambios se pueden atribuir a mejoras en las habilidades del entrevistador. Teniendo más experiencia, los entrevistadores pueden obtener información más completa del entrevistado, resultando en cambios que son diferentes a los efectos de la intervención. Los entrevistadores también pueden descuidarse a medida que se familiarizan con el manual de entrevistas o la lista de verificación.</p> | <ul style="list-style-type: none"> <li>• Mantener la redacción exacta de las preguntas iniciales.</li> <li>• Si se decide medir nuevos indicadores en el seguimiento, agregar preguntas totalmente nuevas al instrumento de encuesta y analizarlas con un diseño solamente con postest.</li> <li>• Recolectar datos acerca de los nuevos indicadores usando un método distinto de recolección de datos (por ejemplo, usar grupos focales o entrevistas finales a los clientes).</li> <li>• Asegurar que los recolectores de datos están bien capacitados antes de la primera ronda de recolección de datos.</li> <li>• Asegurar que los entrevistadores no sepan quienes fueron asignados a los grupos de participantes y de control, para que no anticipen las respuestas.</li> </ul>  |

| Amenaza a la validez  | Cómo minimizar la amenaza  |
|---|--|
| <p>Los <b>cambios en la implementación</b> del programa también pueden afectar la validez de la evaluación. Por ejemplo, un programa que se suspende, intensifica o cambia su contenido puede producir resultados diferentes a los planeados.</p> | <ul style="list-style-type: none"> <li>• Seleccionar intervenciones bien establecidas y que tengan menos probabilidades de ser cambiadas o amenazadas por eventos externos, tales como cambios de políticas o pérdida de financiamiento.</li> <li>• Asegurar que los que están implementando el programa entienden los efectos de cambiar la intervención, y hacerlos documentar y justificar cualquier cambio.</li> <li>• Efectuar una evaluación de proceso para documentar cómo cambió el programa durante el transcurso de la implementación.</li> </ul> |

### *...Elección de un diseño de estudio para un programa en ejecución*

#### *Las opciones para iniciar la evaluación y la capacidad de demostrar el impacto son más limitadas a medida que progresa el programa*

Las opciones de evaluación son limitadas si se está a medio camino o cerca del final del programa o ciclo de programa. Si este es el caso, la evidencia del impacto probablemente no será tan sólida como hubiera sido con una evaluación al comienzo del programa o ciclo de programa. Si hay que seleccionar un diseño de estudio después de comenzado el programa, se debe considerar lo siguiente:

#### **Opciones para elegir un diseño de estudio**

##### **Programa a medio ejecutar**

Si se hicieron mediciones base en los indicadores clave de resultados:

- Continuar el seguimiento de estos indicadores en las recolecciones de datos que se hagan cuando el programa está a la mitad o al final. Esto será igual que un diseño no experimental con pretest-postest. También se debería averiguar si se están efectuando encuestas repetidas en el mismo lugar, las cuales se pueden usar como población de control genérico.

Si no se recolectaron datos base acerca de los indicadores clave de resultados, hay dos opciones:

- Recolectar los datos en ese momento, tratarlos como datos base y efectuar una nueva ronda de recolección de datos cerca del final del programa o ciclo del programa. Esto será nuevamente lo mismo que el diseño no experimental con pretest-postest, pero cubrirá sólo una parte del período del programa. Se perderá la porción del impacto del programa que ocurrió antes de la ronda inicial de recolección de datos, pero si el programa se implementa lentamente, esto puede no ser un problema grave.
- Esperar hasta el final del programa y usar un diseño solamente con postest.

Inclusión de un grupo de control para fortalecer la evidencia del impacto

Como las dos opciones ya descritas son diseños bastante débiles, se debe explorar la posibilidad de incluir un grupo de control o usar un control genérico (si hay alguno disponible en la localidad). Tener algún tipo de población de control fortalece la evidencia acerca del impacto del programa.

## Opciones para elegir un diseño de estudio

### Al final del programa o ciclo del programa

Si se hicieron mediciones base sobre indicadores clave de resultados:

- Medir estos indicadores nuevamente en una ronda final de recolección de datos. Esto será un diseño no experimental con pretest-posttest. También se debe verificar si se han efectuado encuestas repetidas en el lugar, de las cuales la información recabada puede usarse como un control genérico.

Si no se recolectaron datos base acerca de los indicadores clave de resultados, hay varias opciones—algunas más deseables que otras:

- Usar un diseño no experimental solamente con posttest y medir los indicadores al final del período del programa. Sin embargo, este es un diseño muy débil. Las siguientes opciones son preferibles:
- Incluir un grupo de control (este sería un diseño no equivalente de grupo de control solamente con posttest).
- Usar un control genérico (esto sería un diseño de control genérico solamente con posttest).

Implicaciones de la recolección de datos al final del programa

Cualquiera que sea la opción que se elija, esperar hasta el final del programa o ciclo de programa para recolectar datos para medir el impacto significará que:

- se necesitará hacer un análisis estadístico relativamente complicado para poder producir evaluaciones de impacto con resultados creíbles.
- se corre el riesgo de no poder atribuir los cambios en la población objetivo a los esfuerzos del programa.

## Resumen de diseños de estudio para evaluaciones de impacto

| Tipo de diseño de estudio | Característica determinante  | Diseño de estudio específico                                 | Características determinantes  | Ventajas  | Debilidades  | Situaciones en las que se puede usar este diseño   |
|---------------------------|--|--|--|---|--|--|
| Experimentos aleatorios   | Los participantes son asignados aleatoriamente a los grupos de tratamiento y control | <i>Grupo de control con pretest-postest</i>                  | <ul style="list-style-type: none"> <li>Asignación aleatoria</li> <li>Se miden los indicadores al inicio y al final</li> <li>Impacto se mide como la diferencia entre los cambios en los indicadores para los grupos de tratamiento y de control</li> </ul> | Diseño más riguroso, el más alto nivel de validez   | <ul style="list-style-type: none"> <li>La asignación aleatoria no siempre es factible.</li> <li>Se deben mantener condiciones experimentales a través de todo el programa</li> </ul>   | Cuando se está comenzando un nuevo programa o ciclo de programa  |
|                           |  | <i>Grupo de control solamente con postest</i>                | <ul style="list-style-type: none"> <li>Asignación aleatoria</li> <li>Medir los indicadores sólo después</li> <li>Impacto se muestra como la diferencia entre los indicadores para los grupos de intervención y control</li> </ul>                          | No se tiene que recolectar datos base   | <ul style="list-style-type: none"> <li>Más débil que el pretest-postest (sin conocimientos sobre la situación pre-intervención)</li> <li>Se deben mantener condiciones experimentales a través de todo el programa</li> <li>Menos poder estadístico</li> </ul>                               | Cuando se está dispuesto a confiar en que la asignación aleatoria ha creado grupos experimentales equivalentes (para que no se necesiten datos base) |
| Quasi-experimentos        | El grupo de control es similar al grupo de intervención, pero no equivalente.        | <i>Grupo de control no equivalente con pretest-postest.</i>  | Similar al “grupo de control con pretest-postest”, pero los grupos experimentales se eligen haciendo coincidir las características en vez de por asignación al azar.   | Más fácil elegir los grupos por coincidencia de características que por asignación al azar. | <ul style="list-style-type: none"> <li>Se debe mantener las condiciones experimentales a través del programa</li> <li>Riesgo de que los grupos experimentales puedan diferir en factores que la coincidencia de características o el análisis estadístico no puedan contabilizar.</li> </ul> | <ul style="list-style-type: none"> <li>Cuando se han recolectado datos base</li> <li>Cuando la asignación al azar no es factible.</li> </ul>         |
|                           |  | <i>Grupo de control no equivalente solamente con postest</i> | Similar al “grupo de control solamente con postest”, pero los grupos experimentales se eligen haciendo   | No se tiene que recolectar datos base   | <ul style="list-style-type: none"> <li>Se deben mantener las condiciones experimentales a través del programa</li> </ul>   | <ul style="list-style-type: none"> <li>Cuando no se recolectaron datos base</li> </ul>   |

|  |  |                         |   |   |  |   |
|--|--|-------------------------|---|---|--|---|
|  |  |                         | coincidir las características en vez de por asignación al azar                                |   | <ul style="list-style-type: none"> <li>• Alto riesgo que los grupos experimentales puedan diferir en factores que la coincidencia de características o el análisis estadístico no puedan controlar</li> <li>• Menos poder estadístico</li> </ul> | <ul style="list-style-type: none"> <li>• Cuando la asignación al azar no es factible</li> </ul>   |
|  |  | <i>Control genérico</i> | Comparar los cambios en los indicadores del programa con la situación de la población general | No hay necesidad de crear grupos de control | Rara vez hay disponibilidad de datos para los indicadores relevantes y una población para comparación apropiada—hay oportunidades limitadas para usar este diseño  | <ul style="list-style-type: none"> <li>• Cuando los indicadores para el control genérico de la población son relevantes al programa</li> <li>• Cuando la población de control genérico es similar a la población objetivo del programa</li> </ul> |

| Tipos de diseño de estudio | Característica determinante | Diseño de estudio específico | Características determinativas  | Ventajas   | Debilidades   | Situaciones en las que se puede usar este diseño  |
|----------------------------|-----------------------------|------------------------------|---|--|---|---|
| Diseños no experimentales  | Sin grupo de control        | <i>Seriado</i>               | <ul style="list-style-type: none"> <li>• Medir los indicadores varias veces antes y después de la intervención</li> <li>• El impacto se muestra como un cambio en las tendencias después de la intervención</li> </ul>                              | El diseño más sólido entre los no experimentales, si se tiene una serie suficientemente larga antes de la intervención                         | Se deben haber efectuado por lo menos tres o más observaciones de preintervención.  | <ul style="list-style-type: none"> <li>• Cuando no se ha planificado un grupo de control</li> <li>• Cuando se está evaluando un programa de cobertura total</li> </ul>  |
|                            |                             | <i>Pretest-postest</i>       | <ul style="list-style-type: none"> <li>• Medir los indicadores para la población objetivo del programa antes y después</li> <li>• Hacer pruebas estadísticas para cualquier cambio observado para verificar su significación estadística</li> </ul> | Se pueden documentar los cambios en los indicadores de resultados del programa en la población objetivo  | <ul style="list-style-type: none"> <li>• Difícil atribuir el impacto al programa</li> <li>• Requiere de análisis sofisticado para fortalecer los hallazgos</li> </ul> | <ul style="list-style-type: none"> <li>• Igual que anterior</li> <li>• Cuando no se tienen suficientes datos de antes de la intervención para un diseño seriado</li> </ul>  |
|                            |                             | <i>Solamente postest</i>     | <ul style="list-style-type: none"> <li>• Medir los indicadores para la población objetivo del programa sólo después de la intervención</li> </ul>   | Ninguna para medir el impacto—los datos se pueden usar para informar un nuevo diseño de programa o como base para el próximo ciclo de programa | El diseño más débil no mide ni los cambios en los indicadores a través del tiempo, ni las diferencias con un grupo de control   | <ul style="list-style-type: none"> <li>• Igual que para el diseño seriado</li> <li>• Cuando no se tiene datos de pre-intervención</li> <li>• Cuando se necesitan datos descriptivos de post-intervención</li> </ul> |

## □ Capítulo 6: Muestreo

Una mirada al capítulo:

- *Describe los tipos de métodos de muestreo y las formas de determinar cuál método de muestreo es apropiado para el programa*
- *Se centra en un método de uso común de muestreo, el muestreo por conglomerados*
- *Examina cómo determinar y calcular el tamaño de la muestra que se necesita para actividades de recolección de información*

### *...¿Qué es el muestreo y qué papel juega en la evaluación de un programa?*

Se pueden usar informes de programas, estadísticas de servicios y datos recolectados de manera habitual para medir algunos indicadores como parte de la evaluación. Sin embargo, en algunos casos se necesitará recolectar datos que van más allá de los registros existentes. Por ejemplo, si se quieren medir los cambios en los conocimientos sobre el SIDA entre los jóvenes a los que llega el programa. Puede ser imposible recolectar datos de todos los jóvenes involucrados en el programa dado su gran número y los recursos limitados del programa. Sin embargo, sería posible recolectar datos de una muestra o subconjunto de jóvenes. El muestreo permite recolectar una menor cantidad de datos que representan a todo el grupo. Esto ahorrará tiempo, dinero y otros recursos sin comprometer la confiabilidad de la información.

El muestreo es el proceso de elegir sistemáticamente un subconjunto del total de la población que interesa encuestar. Con el muestreo, se obtiene información que se puede generalizar a toda la población objetivo del programa.

El muestreo es útil tanto para las evaluaciones de proceso como para las de resultados. Por ejemplo, al efectuar evaluaciones de proceso se puede:

- *elegir una muestra de las instalaciones para evaluar si tienen los equipos e insumos necesarios para implementar el programa apropiadamente*
- *elegir una muestra de transacciones de servicio de salud para evaluar si los proveedores de servicio están siguiendo las pautas de servicio*
- *elegir una muestra de hogares o comunidades para evaluar la cobertura del programa*

Para las evaluaciones de resultados o impactos, el muestreo permite:

- *elegir una muestra de jóvenes en el programa o en la población general para evaluar sus niveles de comportamientos arriesgados y/o*
- *elegir una muestra de escuelas para evaluar si las habilidades para la vida han mejorado como resultado del programa*

## Términos clave del muestreo

**Muestra:** Una parte del todo seleccionada para representarlo.

**Elemento:** La persona de la cual se recolectarán datos; un elemento puede ser un joven, un padre o un proveedor de servicios.

**Conglomerado:** Un grupo agregado de elementos de los cuales se recolectarán datos; un conglomerado puede ser una sala de clases, una escuela, un servicio de salud o un grupo juvenil.

**Marco de muestreo:** Una lista global de todos los elementos relevantes o conglomerados que se usan para seleccionar la muestra.

### Se puede buscar ayuda externa

Cuando el muestreo se hace bien, es un aspecto clave de la credibilidad de una evaluación. Sin embargo, el muestreo puede ser complicado, especialmente cuando se evalúa un programa grande y complejo.

Después de leer este capítulo, se puede buscar asistencia técnica para asegurar que el esquema de muestreo desarrollado es sólido en términos estadísticos. Expertos de las universidades locales e institutos de investigación pueden ayudar a desarrollar un esquema de muestreo eficiente para la evaluación del programa.

### Los métodos de muestreo se dividen en dos categorías amplias: probabilístico y no probabilístico

Se necesitará elegir el método de muestreo. Los métodos de muestreo probabilístico fortalecen la validez de los resultados de la evaluación. Sin embargo, los métodos de muestreo probabilísticos pueden ser lentos, costosos y requieren un cierto nivel de conocimientos para ejecutarlos. Los métodos de muestreo no probabilísticos son más flexibles, menos costosos y requieren menos tiempo, pero los resultados de la evaluación son más débiles.

Los métodos de muestreo probabilístico están basados en la teoría de la probabilidad, un concepto matemático que se refiere a la capacidad de predecir la probabilidad estadística de que un evento cualquiera pueda ocurrir. Los métodos de muestreo probabilístico requieren que:

- *todos los elementos del programa tengan alguna posibilidad de ser seleccionados (nota: los elementos no tienen que tener la misma probabilidad, basta con que sea una posibilidad)*
- *sea posible calcular la probabilidad de que cada elemento pueda ser elegido*
- *el azar determine los elementos que son escogidos.*

A la inversa, los métodos de muestreo no probabilísticos son enfoques que no están basados en la teoría de la probabilidad. Con estos métodos, las muestras no son seleccionadas aleatoriamente. Estos métodos incluyen:

- *obtener una muestra de sujetos en base a la oportunidad*
- *usar el buen juicio para elegir lo que parece ser una muestra representativa de la población más amplia*
- *seleccionar nuevos sujetos si son referidos por otras personas que ya han sido seleccionadas para el grupo de estudio*

El cuadro que sigue describe métodos de muestreo específicos que corresponden a las dos categorías.



## Tipos de métodos de muestreo

| Métodos de muestreo probabilístico  | Métodos de muestreo no probabilístico   |
|---|---|
| <p><b>Muestreo aleatorio simple</b> - Los elementos se eligen aleatoriamente y cada uno tiene la misma probabilidad de ser elegido. Por ejemplo, los elementos se pueden seleccionar sacándolos de un sombrero o, idealmente, de una tabla con números aleatorios en un texto de estadísticas o generados por un computador.</p> <p><b>Muestreo sistemático</b> - El primer elemento se elige aleatoriamente. Los siguientes se eligen usando un intervalo fijo (por ejemplo, cada décimo elemento) hasta que se obtenga el tamaño deseado de la muestra.</p> <p><b>Muestreo estratificado</b> - La población de la que se obtendrá la muestra se divide en grupos homogéneos basados en características que se consideran importantes para los indicadores que serán medidos, tales como jóvenes sexualmente activos. Entonces se elige un muestreo simple o sistemático de cada grupo.</p> <p><b>Muestreo de conglomerados</b> - En primer lugar, se elige una muestra aleatoria simple de conglomerados de un marco de muestreo. Los conglomerados incluyen escuelas, servicios de salud o clubes juveniles. Después, se selecciona una muestra aleatoria simple dentro de cada conglomerado. Los individuos así seleccionados conforman la muestra.</p> <p><b>Muestreo de etapas múltiples</b> - Similar al muestreo de conglomerados, pero con varias etapas de muestreo y submuestreo. Este método usualmente se usa en encuestas de población a gran escala.</p> | <p><b>Muestreo por conveniencia</b> - Se saca una muestra en base a la oportunidad. Por ejemplo, la muestra puede incluir a jóvenes que estén presentes en una actividad de la escuela, proveedores de servicios que estén asistiendo a una conferencia o padres asistiendo a un evento en la escuela.</p> <p><b>Muestreo por cuotas</b> - Un marco de muestreo se define antes de recolectar los datos y se selecciona una muestra de esta lista, pero no al azar.</p> <p><b>Muestreo en cascada</b> - Se recolectan datos de un pequeño grupo de personas con características especiales, a los cuales se les pide que identifiquen a otras personas como ellos. Se recolectan datos de estas referencias, a los cuales también se les pide que identifiquen a personas como ellos. Este proceso continúa hasta que se haya logrado el tamaño objetivo de la muestra o hasta que la recolección adicional de datos no proporcione más información nueva. Este método también se conoce como muestreo de redes o muestras en cadena.</p> |

### ...¿Qué método de muestreo es mejor?

Para seleccionar el mejor método de muestreo para las necesidades específicas del programa, se deben considerar los siguientes factores:

- *tiempo y recursos disponibles*
- *tamaño del programa*
- *indicadores que se planea medir y*
- *fuentes de datos de donde se planea recolectar información*

El cuadro siguiente resume las principales ventajas y desventajas de los métodos de muestreo probabilísticos y no probabilísticos.

## Ventajas y desventajas de los muestreos probabilísticos y no probabilísticos

| Tipo de muestreo           | Ventajas  | Desventajas  |
|----------------------------|---|--|
| Muestreo probabilístico    | <ul style="list-style-type: none"> <li>• Menos propenso al sesgo</li> <li>• Permite una estimación de la magnitud del error de muestreo, de la cual se puede determinar la importancia estadística de los cambios/diferencias en los indicadores</li> </ul> | <ul style="list-style-type: none"> <li>• Requiere una lista de todos los elementos de la muestra</li> <li>• Más lenta</li> <li>• Más costosa</li> <li>• No es ventajosa cuando se elegirá un número pequeño de elementos</li> </ul>  |
| Muestreo no probabilístico | <ul style="list-style-type: none"> <li>• Más flexible</li> <li>• Menos costoso</li> <li>• Requiere menos tiempo</li> <li>• Se pueden preferir muestras seleccionadas según el criterio del analista cuando se use un pequeño número de elementos</li> </ul> | <ul style="list-style-type: none"> <li>• Mayor riesgo de sesgo</li> <li>• Puede no ser posible generalizar los resultados a la población objetivo del programa</li> <li>• La subjetividad puede hacer difícil la medición de los cambios en los indicadores a través del tiempo.</li> <li>• No hay forma de evaluar la precisión o confiabilidad de los datos</li> </ul> |

### **El muestreo probabilístico es el preferido para la mayoría de los tipos de evaluaciones de programas juveniles**

Si se está trabajando con un programa que tiene una población objetivo de más de 100 jóvenes y se quiere medir cambios en los indicadores relacionados a los objetivos del programa, los métodos de muestreo probabilístico darán resultados de evaluación con más objetivos y defendibles desde un punto de vista científico. Los métodos de muestreo no probabilístico, a pesar de ser menos costosos y menos lentos, pueden producir cierta vulnerabilidad acerca de si los resultados de la evaluación son representativos y sin sesgos. La razón principal por la que algunos programas no usan muestreo probabilístico para sus evaluaciones es la falta de tiempo, recursos o conocimientos especializados para efectuar este tipo de muestreo.

### **El muestreo no probabilístico debería ser usado en algunas circunstancias**

Existen algunas circunstancias en las que el muestreo probabilístico puede no ser posible o preferible:

- **Evaluaciones de programas más pequeños de jóvenes.** Cuando se está eligiendo un pequeño número de elementos (menos de 20), no son aplicables las propiedades estadísticas clave del muestreo probabilístico. Si se planea recolectar datos de menos de 20 individuos, una muestra no probabilística por cuotas que se considere representativa en términos de características importantes es tan efectiva como el muestreo probabilístico.
- **La evaluación de programas cuya población objetivo es difícil de definir.** Algunos programas trabajan con subgrupos de jóvenes, tales como niños de la calle

o consumidores de drogas. En estas situaciones, puede ser difícil o imposible construir un marco de muestreo o una lista de todos los jóvenes que pertenecen a la población objetivo. Sin un marco de muestreo no se puede efectuar un muestreo probabilístico. En estos casos la opción más factible puede ser el muestreo en cascada.

- ***Cuando se efectúan discusiones de grupos focales.*** Es difícil tratar de reunir a un grupo de personas elegido aleatoriamente en un lugar dado para tener un debate de grupo focal. Por lo tanto, el muestreo por cuotas o de conveniencia se usa a menudo. De estos, el muestreo por cuotas es el preferido porque es menos propenso al sesgo.

### ***... Muestreo por conglomerados***

El muestreo por conglomerados es el tipo de muestreo probabilístico más usado y es útil para la mayor parte de las evaluaciones de programas de SSRA.

#### ***El muestreo de conglomerados involucra dos etapas***

En primer lugar, se identifican los grupos de los cuales se quieren recolectar datos. Estos grupos son los conglomerados. Ejemplos de conglomerados son las escuelas, servicios de salud o clubes juveniles. Se desarrolla un marco de muestreo, o lista de todos los conglomerados, y se selecciona una muestra de conglomerados de esta lista.

En segundo lugar, se desarrolla una lista de elementos de muestra dentro de los conglomerados de muestreo seleccionados. Ejemplos de elementos son estudiantes, clientes de los servicios de salud o miembros de clubes juveniles. Luego se selecciona una muestra de elementos de cada conglomerado. El resultado es la muestra de individuos de los cuales se recolectarán datos.

Por ejemplo, si una organización tiene un programa de adolescentes educadores en 40 escuelas, cada escuela sería considerada un conglomerado. Primero se selecciona una muestra de escuelas entre todas las que participan en el programa. Luego se selecciona una muestra aleatoria de alumnos de cada escuela. Esta sería la muestra de la cual se recolectarán los datos.

El muestreo de conglomerados es el tipo de muestreo probabilístico menos costoso, porque no requiere que se desarrolle un marco de muestreo de todos los elementos antes de seleccionar una muestra. Desarrollar marcos de muestreo individuales puede ser costoso. Por ejemplo, puede ser difícil desarrollar una lista de todos los clientes que buscaron servicios en las 12 clínicas de las cuales se quieren tomar muestras. Sería más fácil listar todas las clínicas. Cada clínica será un conglomerado. Después de seleccionar una muestra de clínicas, sólo queda desarrollar listas de clientes de aquellas clínicas incluidas en la muestra final.

A continuación se presenta un procedimiento general, paso a paso, para hacer muestreos por conglomerados una vez que se han identificado la población objetivo y los indicadores que se usarán en la evaluación.

- **Primer paso: Definir los "conglomerados" que se usarán**

Un conglomerado es un grupo claramente definido de elementos de muestreo de los cuales se puede seleccionar una submuestra más pequeña. Algunos ejemplos de conglomerados incluyen:

- *áreas geográficas con fronteras fijas (para las encuestas de hogares)—los hogares e individuos que responden son las submuestras que se elegirán en las etapas siguientes de selección*
- *escuelas*
- *servicios de salud y*
- *clubes juveniles*

Por ejemplo, las clínicas pueden servir de conglomerados para medir los indicadores de proceso y de resultado. Para medir el impacto, las áreas geográficas preparadas por una oficina nacional de estadísticas (áreas de enumeración censal o AEC) se pueden usar como conglomerados.

La elección de qué unidad usar como conglomerado dependerá de los datos disponibles. Es importante recordar que con el muestreo de conglomerados se puede minimizar el esfuerzo necesario para desarrollar un marco de muestreo. Por lo tanto, la unidad que se usará de conglomerado debería depender de los grupos en que están involucrados los elementos que interesan para la muestra. El recuadro a continuación ilustra posibles conglomerados para una variedad de elementos posibles de una muestra.

| <b>Elementos de la muestra y posibles conglomerados</b> |  |
|---|--|
| <b>Elemento de la muestra</b>                           | <b>Posibles conglomerados</b>  |
| Jóvenes   | <ul style="list-style-type: none"> <li>• Áreas geográficas/comunidades/hogares</li> <li>• Escuelas</li> <li>• Cursos/ secciones del colegio</li> <li>• Organizaciones juveniles/clubes deportivos</li> <li>• Servicios de salud</li> <li>• Lugares de trabajo</li> <li>• Jóvenes nombrados como contactos por un adolescente educador</li> </ul> |
| Padres (y miembros de la familia extendida)             | <ul style="list-style-type: none"> <li>• Vecindarios, aldeas u otras áreas geográficas</li> <li>• Aquellos cuyos hijos jóvenes han sido seleccionados para la muestra (usando los mismos conglomerados de más arriba)</li> </ul>   |
| Profesores  | <ul style="list-style-type: none"> <li>• Escuelas</li> <li>• Cursos/secciones en la escuela</li> </ul>   |
| <b>Elemento de la muestra</b>                           | <b>Posibles conglomerados</b>  |
| Líderes comunitarios                                    | <ul style="list-style-type: none"> <li>• Áreas geográficas/comunidades</li> <li>• Instituciones religiosas</li> <li>• Instituciones comerciales</li> <li>• Oficinas gubernamentales</li> </ul>   |
| Proveedores de servicios                                | <ul style="list-style-type: none"> <li>• Hospitales/clínicas</li> <li>• Líneas telefónicas de apoyo</li> <li>• Visitas al hogar</li> </ul>   |
| Eventos de los servicios de salud                       | <ul style="list-style-type: none"> <li>• Hospitales/clínicas</li> <li>• Farmacias</li> <li>• Turnos de trabajo dentro de un servicio de salud</li> </ul>   |
| Educadores pares  | <ul style="list-style-type: none"> <li>• Escuelas</li> <li>• Organizaciones juveniles</li> <li>• Lugares de trabajo</li> </ul>   |
| Trabajadoras sexuales                                   | <ul style="list-style-type: none"> <li>• Prostíbulos</li> <li>• Bares/clubes</li> <li>• Rutas comerciales</li> </ul>   |

- **Segundo paso: Desarrollo del marco de muestreo**

En el muestreo por conglomerados, el marco del muestreo será una lista de conglomerados. En algunos casos, la lista de conglomerados que se ha definido puede estar disponible desde antes, aunque se debe asegurar que la lista esté completa y actualizada, o se tendrá que crear un nuevo marco de muestreo. Para crear el nuevo marco, puede que se necesite visitar las comunidades dentro del área objetivo del programa para hacer una lista de escuelas, servicios de salud u organizaciones juveniles. Tal vez se necesite preparar mapas simplificados de las aldeas, manzanas de la ciudad o vecindarios para usarlos como conglomerados.

La mayoría de los programas de salud tienen una lista de clínicas afiliadas. Si se está usando las Áreas de Enumeración Censales (AEC) de la oficina nacional de estadísticas como conglomerados para efectuar encuestas de hogares sobre los jóvenes, esa lista debería estar disponible. Si no, deberá efectuarse un trabajo preliminar en terreno para desarrollar una lista de manzanas o vecindarios.

Si crear una lista de todos los elementos dentro de la población objetivo no es factible financiera o logísticamente, se puede restringir la muestra a una parte de la población objetivo. Sin embargo, los hallazgos de la evaluación sólo serán válidos para la parte de la población objetivo incluida en el marco de muestreo. Por ejemplo, puede que sea necesario, por razones logísticas, incluir sólo las clínicas ubicadas en la parte central de la ciudad. Una muestra seleccionada de este marco restringido podría dar como resultado hallazgos de la evaluación que sólo corresponden a las clínicas de la parte central de la ciudad. Las conclusiones acerca de los efectos del programa para toda la ciudad sólo podrían hacerse bajo el supuesto de que otras clínicas fueron igualmente efectivas que las del área central de la ciudad.

### **Eliminación del sesgo**

El sesgo es el grado en que el estimado de un indicador encontrado por una encuesta es diferente de su valor "real". Para evitar el sesgo, los marcos de muestreo deben cubrir a toda la población objetivo en la evaluación de un programa. Por ejemplo, si sólo se incluyeron a los participantes activos en el programa, puede obtenerse una visión sesgada de las fortalezas del programa. Al no incluir en la muestra a los jóvenes que desertaron el programa, por ejemplo, se perdería información sobre las razones por las que a algunos jóvenes no les gustaron las intervenciones.

- **Tercer paso: Seleccionar una muestra de conglomerados**

Una vez que se ha desarrollado el marco de muestreo, el próximo paso es seleccionar una muestra de conglomerados usando el muestreo aleatorio *simple* o el muestreo *sistemático*. Por ejemplo, para la encuesta de hogares, se puede decidir usar una muestra de AEC usando el muestreo sistemático. Los pasos específicos involucrados en el muestreo sistemático se describen en el Anexo 1.

Determinar el mejor método dependerá del número de conglomerados en el marco de muestreo y de algunas consideraciones estadísticas tratadas más adelante en este capítulo y en el Anexo 1. La pregunta de cuántos conglomerados se deben elegir también se trata en el Anexo 1.

- **Cuarto paso: Seleccionar una muestra de elementos de los conglomerados elegidos**

El próximo paso será elegir una muestra de elementos de los conglomerados que se han seleccionado. A menudo esto se hace usando el muestreo aleatorio simple o el muestreo sistemático. Sin embargo, se pueden usar otras técnicas de muestreo dependiendo del contexto. Por ejemplo, después de seleccionar una muestra de clínicas, se pueden efectuar entrevistas de salida a todos los clientes jóvenes que llegan a las clínicas de la muestra en un

día elegido al azar. Las pautas para decidir qué métodos de muestreo usar en distintos escenarios se dan en el Anexo 1. El recuadro que aparece a continuación resume cómo aplicar los pasos básicos en el muestreo por conglomerados a las encuestas de hogares, escuelas y servicios de salud.

| Esquemas de muestreo de conglomerados para encuestas de hogares, escuelas y servicios de salud |                      |  |   |  |  |   |  |
|--|----------------------|--|---|--|--|---|--|
| Tipo de encuesta   | Población objetivo   | Definición de conglomerado                                       | Marco de muestreo (conglomerados)                       | Seleccionar muestras de conglomerados  | Marco de muestreo (elementos de la muestra)  | Seleccionar una muestra de elementos de la muestra  | Para más explicaciones ver:                              |
| Hogares  | Jóvenes, padres      | Área geográfica con fronteras fijas: AEC, manzanas, aldeas, etc. | Lista de AECs o de otras áreas geográficas              | Usar muestreo sistemático para seleccionar las AEC (u otras áreas geográficas) de las cuales se tomarán muestras | Lista de hogares en un conglomerado seleccionado   | Usar el muestreo aleatorio simple o el sistemático para elegir un número igual de hogares de cada conglomerado de la muestra  | Anexo 1 para métodos alternativos de muestreo<br>Anexo 2 |
|  | Líderes comunitarios | Igual que el anterior  | Igual que el anterior                                   | Igual que el anterior  | Lista de líderes de la comunidad en un conglomerado seleccionado                           | Usar el muestreo aleatorio simple para elegir un número igual de líderes de cada conglomerado de la muestra, o seleccionar una muestra de conglomerados e incluir a todos los líderes de la comunidad de esos conglomerados |  |
| Escuelas   | Estudiantes y padres | Escuelas   | Lista de todas las escuelas a las que llega el programa | Usar muestreo sistemático para seleccionar las escuelas en las cuales se hará el muestreo                        | Lista de todos los estudiantes en cada escuela o lista de todos los cursos en cada escuela | Usar muestreo aleatorio simple o sistemático para elegir las muestras de estudiantes de una lista de todos los estudiantes, o elegir muestras de cursos e incluir a todos los estudiantes                                   | Anexo 2  |

| Tipo de encuesta  | Población objetivo           | Definición de conglomerado     | Marco de muestreo (conglomerados)                       | Seleccionar muestras de conglomerados  | Marco de muestreo (elementos de la muestra)  | Seleccionar una muestra de elementos de la muestra   | Para más explicaciones ver: |
|-------------------|------------------------------|--------------------------------|---|--|--|--|-----------------------------|
| Escuelas (cont.)  | Profesores y administradores | Igual que más arriba           | Igual que más arriba                                    | Igual que más arriba   | Lista de todos los profesores en cada escuela, o lista de los cursos en cada escuela | Usar muestreo aleatorio simple o sistemático para elegir profesores de una lista de todos los profesores, o elegir al profesor asociado con la muestra de clases elegidas para las encuestas de estudiantes (estrategia de vinculación)        |                             |
| Servicio de salud | Proveedores de servicios     | Unidades de servicios de salud | Lista de todas las unidades a las que llega el programa | Usar muestreo sistemático para seleccionar las unidades de las cuales se tomará la muestra | Lista de proveedores de servicios en cada unidad de la muestra                       | Usar muestreo aleatorio simple o sistemático para elegir proveedores de una lista de todos los proveedores en cada unidad, o entrevistar a todos los proveedores presentes en un día elegido al azar en que se visite la unidad de la muestra. |                             |

### *...Temas clave en el muestreo por conglomerados*

Al elegir las muestras, se debe tener en mente ciertos temas y consideraciones:

- *Decidir si tomar o no muestras de conglomerados o elementos, o si incluirlos a todos*
- *Determinar cuántos conglomerados elegir en una muestra*
- *Cómo utilizar el muestreo por conglomerados con conglomerados de diferentes tamaños*
- *Elegir distintos elementos de la muestra para la misma encuesta*
- *Muestreo para subgrupos y*
- *Decidir cuándo usar encuestas nacionales de hogares para evaluar programas basados en escuelas o en unidades de servicios de salud*

### **Decidir si tomar o no tomar muestras por conglomerados o elementos, o si incluirlos todos**

Hay algunas situaciones en las que no se necesita tomar muestras por conglomerados o elementos. Decidir si tomar muestras o no depende de:

- **cuántos conglomerados se tengan:** Para programas de pequeña escala que operan sólo en unas pocas localidades (por ejemplo, comunidades, escuelas o servicios), se puede incluir a todas las localidades en la evaluación.
- **cuántos elementos existen dentro de cada conglomerado:** Si el número de elementos en cada localidad no es muy grande, logísticamente puede ser más simple incluir a todos los elementos en el muestreo por conglomerados. Por ejemplo, si se utilizaran cuestionarios autoadministrados para evaluar un programa basado en la escuela (y por supuesto si los recursos lo permiten), sería más fácil incluir a todas las clases en las escuelas de la muestra o a todos los estudiantes en las clases de la muestra. Esto también es cierto para:
  - *administradores y profesores en las escuelas de la muestra*
  - *líderes de la comunidad*
  - *administradores y proveedores de servicios de salud*

### **Determinar cuántos conglomerados se necesitan en una muestra**

Cuando se eligen conglomerados para una muestra, se deben seleccionar el máximo de conglomerados (por ejemplo, comunidades, escuelas, servicios) que permitan los recursos. Generalmente deberían elegirse por lo menos 30 conglomerados, especialmente en el caso de programas grandes. Muestras con más conglomerados de menor tamaño son siempre preferibles a aquellas con menos conglomerados de mayor tamaño. Esto se debe a que elegir un número menor de conglomerados reducirá la precisión o confiabilidad de los datos, haciendo más difícil la detección de verdaderos cambios en los indicadores. Por ejemplo, es mejor elegir 30 escuelas con una población de 200 estudiantes cada una que elegir 20 escuelas con una población de 300 estudiantes cada una.

Si se eligen menos de 20 conglomerados, habrá pocas ventajas en la selección aleatoria de los conglomerados. Esto se debe a que las propiedades estadísticas deseables del muestreo probabilístico no se aplican a muestras con menos de 20 elementos. En estos casos será igualmente aceptable seleccionar una muestra de conglomerados que el evaluador considere “representativos” de la población objetivo. Es decir, el perfil de los conglomerados respecto a características importantes (por ejemplo, tamaño, ubicación, nivel socioeconómico de los alumnos o área de captación servida) debe ser similar a la distribución de esas características, la cual fue observada en la población objetivo en su conjunto. Para lograr esto en la selección de una muestra de escuelas, se debe estar seguro de que hay representación de escuelas grandes y pequeñas, escuelas que sirven a los ricos y a los pobres, y escuelas de diferentes áreas geográficas dentro del área del estudio. En el Anexo 1 se pueden encontrar más criterios para decidir cuántos grupos escoger.

### **Muestreo por conglomerados de diferentes tamaños**

Al hacer muestreos por conglomerados con distintos tamaños (por ejemplo, escuelas con diferente números de alumnos, áreas geográficas con número de habitantes variables), hay que usar el muestreo sistemático con probabilidad proporcional al tamaño (PPT).

La PPT permite darle a los conglomerados más grandes una mayor probabilidad de selección que a los más chicos. Esto compensa por el hecho de que los elementos en los conglomerados más grandes tienen menos posibilidades de ser elegidos en la segunda etapa de la selección.

Para efectuar un muestreo sistemático con PPT, se necesita una **medida de tamaño (MDT)** de los conglomerados. La MDT es un conteo o estimado del número de elementos de la



muestra asociados con cada conglomerado (ver ejemplos más adelante). No se necesitan conteos exactos; las aproximaciones a la MDT son suficientes. Si no hay MDT disponibles, algunos conglomerados pueden ser elegidos usando el muestreo sistemático con probabilidades iguales.

Los ejemplos de MDT para cada tipo de encuesta incluyen:

- **Hogares:** número de hogares, población total estimada, número estimado de jóvenes en el área geográfica.
- **Escuela:** número de estudiantes
- **Servicio de salud:** número de clientes jóvenes que se atienden en el servicio de salud

En el Anexo 1 se pueden encontrar instrucciones detalladas para seleccionar una muestra de conglomerados usando el muestreo sistemático con PPT y con probabilidades iguales.

### **Elegir distintos elementos de muestra para la misma encuesta**

Cuando se están seleccionando varios elementos (por ejemplo, estudiantes, padres, profesores, líderes de la comunidad) para la misma encuesta, se debe usar una de dos formas de selección de esos elementos: independiente o vinculada.

Por ejemplo, en una encuesta de escuelas que incluye a estudiantes, padres y profesores como elementos de esa muestra, hay dos opciones de selección:

- *Elegir los elementos **independientemente** uno del otro. Esto **requiere** un proceso de muestreo separado para **cada elemento**: los estudiantes, profesores y padres.*
- *Incluir a todos los elementos en la misma muestra, **seleccionando** a los padres y profesores asociados con la misma escuela o curso elegidos para la encuesta de estudiantes. Esta **estrategia es la del muestreo vinculado**.*

La ventaja principal de la estrategia de muestreo vinculado es que limitará la recolección de datos al mismo conjunto de grupos, lo que resulta en menores costos de recolección de datos.

La estrategia de muestreo vinculado también permite asociar o relacionar los indicadores de diferentes elementos entre sí durante el análisis. Por ejemplo, si los profesores y padres elegidos para una evaluación de programa son los mismos profesores y padres de los estudiantes en la encuesta, se puede determinar si los cambios en los indicadores de los estudiantes están asociados con las actitudes y comportamientos de sus profesores y padres. (Nota: Las mismas opciones son válidas para la selección de elementos de otras muestras, tales como hogares, jóvenes, padres, líderes de la comunidad (ver Anexo 1 para obtener ejemplos ilustrativos).

### **Muestreo para subgrupos**

Si se quiere medir indicadores para subgrupos dentro de la población objetivo (por ejemplo subgrupos definidos por su género, edad, grupo étnico, situación económica, residencia urbana/rural/periurbana), se tiene que elegir una muestra separada para cada subgrupo, con suficiente número de miembros por grupo, como para llevar a cabo los análisis deseados.

Por ejemplo, en el caso de que se quiera saber si el programa ha sido igualmente efectivo para jóvenes hombres y mujeres, la forma más segura de tener suficientes sujetos

masculinos y femeninos es seleccionar muestras separadas. El proceso de dividir la población objetivo en subgrupos y después tomar una muestra de cada uno se conoce como *estratificación*.

### **Cuándo usar encuestas de hogares en la población general**

Las encuestas de hogares de población general se usan para evaluar los programas basados en escuelas o en unidades de servicio sólo si la proporción de la población servida por el programa representa, por lo menos, un 50% y preferiblemente 75% de la población total.

Las encuestas de hogares de población general no son una forma eficiente de evaluar programas que cubren menos del 50% de la población de jóvenes. La razón es que tendrán que contactarse un gran número de hogares para encontrar un número suficiente de jóvenes que fueran expuestos al programa.

### ***...Determinación del tamaño de la muestra***

Muchos indicadores de resultados de SSRA se miden a nivel de personas o eventos individuales, por lo que los elementos de la muestra puedan incluir jóvenes, padres, líderes de la comunidad, profesores, proveedores de servicios de salud, adolescentes educadores y transacciones de servicios. El **tamaño de la muestra** se refiere al número de elementos necesarios en la muestra para que los hallazgos de la evaluación sean estadísticamente significativos. El tamaño de la muestra depende de un número de factores, incluyendo:

- *los indicadores que se elijan,*
- *la prevalencia de los indicadores en la población en estudio, y*
- *la magnitud de cambio que se quiere medir con precisión.*

Cuando se calcula el tamaño de la muestra, se determina cuántos individuos o eventos se necesita encuestar en cada ronda de recolección de datos para que se puedan medir con precisión los cambios en los indicadores especificados. Esta sección se refiere al cálculo de los tamaños de las muestras para medir:

- *cambios en los indicadores que involucran un solo comportamiento*
- *cambios en los indicadores que involucran más de un comportamiento y*
- *cambios en los indicadores para unidades agregadas, tales como escuelas, unidades de servicio y comunidades*

Para comenzar, se debe tener en cuenta tres puntos acerca del tamaño de la muestra:

- ***Medir los cambios en indicadores seleccionados a través del tiempo requiere un tamaño más grande de muestra.*** El tamaño de esta clase de muestra a lo largo de un tiempo es mayor que el tamaño necesario para medir el indicador en un momento dado.
- ***Esta sección asume que los indicadores se medirán en proporciones.*** La mayoría de las evaluaciones de programas sobre jóvenes usan indicadores que se miden en proporciones.
- ***Esta sección no distingue entre el uso de métodos de muestreo probabilístico y no-probabilístico.*** Sin embargo, cabe notar que tener un tamaño de muestra adecuado no minimizará los riesgos de distorsión en muestras no probabilísticas.

Calcular el tamaño de la muestra para medir los cambios en indicadores que involucran un comportamiento depende de los siguientes cinco factores:

- El **nivel inicial** del indicador (es decir, el nivel del indicador durante la encuesta base). A esto le llamaremos  $P_1$ .
- La **magnitud del cambio** de un indicador que se quiere detectar con precisión. Lo llamaremos  $(P_2 - P_1)$  donde  $P_2$  es el nivel de un indicador en la segunda ronda de recolección de datos.
- La probabilidad con la que se quiere estar seguro que la magnitud del cambio  $(P_2 - P_1)$  no ocurrió por casualidad. Esto se llama **nivel de significancia**.
- La probabilidad con que se quiere estar seguro de detectar la magnitud del cambio  $(P_2 - P_1)$  si este ocurrió. A esto se le llama **poder**.
- La proporción de elementos en la muestra de la población objetivo que tiene las características especificadas en el indicador a medir.

Para ayudar a determinar cuán grande debe ser la muestra para medir cambios en indicadores de un comportamiento, en el recuadro a continuación se indican tamaños de muestra para diferentes magnitudes de cambio en los indicadores. Los tamaños de muestra listados se basan en un 90% de nivel de significancia y poder.

| Tamaño mínimo de muestra para medir cambios |   |                                 |
|---|---|---------------------------------|
| Nivel inicial del indicador ( $P_1$ )       | Nivel de indicador que se espera lograr ( $P_2$ ) | Tamaño de muestra necesario (n) |
| 0,10  | 0,20  | 438                             |
| 0,10  | 0,25  | 216                             |
| 0,20  | 0,30  | 638                             |
| 0,20  | 0,35  | 300                             |
| 0,30  | 0,40  | 775                             |
| 0,30  | 0,45  | 353                             |
| 0,40  | 0,50  | 843                             |
| 0,40  | 0,55  | 376                             |
| 0,50  | 0,60  | 843                             |
| 0,50  | 0,65  | 368                             |
| 0,60  | 0,70  | 775                             |
| 0,60  | 0,75  | 330                             |
| 0,70  | 0,80  | 638                             |
| 0,70  | 0,85  | 267                             |
| 0,80  | 0,90  | 438                             |
| 0,80  | 0,95  | 163                             |

Los tamaños de la muestra estiman un nivel de significancia y poder de 90%, un efecto de diseño de 2,0 y pruebas estadísticas de dos colas. Para leer la explicación de estos parámetros, ver Anexo 2.

Para usar el cuadro:

- **Elegir el nivel inicial de un indicador ( $P_1$ ).** Por ejemplo, el indicador puede ser la proporción de jóvenes que pretenden usar preservativos en futuros encuentros sexuales, con el objetivo de aumentar esa proporción. En el momento de la primera encuesta, la mejor estimación es que alrededor del 30% de los jóvenes pretenden usar preservativos. Por lo tanto,  $P_1$  se establece como 0,30.

*Nota:* Si no es posible estimar el nivel inicial de un indicador, el curso más seguro a seguir es elegir un valor para  $P_1$  igual a 0,50, porque esto asegurará un tamaño adecuado de muestra aún si el valor real de  $P_1$  es diferente a 0,50.

- **Especificar la magnitud del cambio que se quiere detectar con precisión ( $P_2 - P_1$ ).** Por ejemplo, suponiendo que el programa fuera considerado un éxito si la proporción de jóvenes que pretende usar preservativos aumentara 15 puntos porcentuales (a 45%). Así,  $P_2$  se establece en 0,45.

*Nota:* Sólo se permiten cambios de 10 a 15 puntos porcentuales en el cuadro. Medir cambios más pequeños requiere tamaños de muestra más grandes, porque se necesita mayor precisión (y por ende más sujetos) para medir cambios pequeños. Los cambios menores a 10 puntos porcentuales requieren un tamaño de muestra que puede exceder los recursos disponibles para la evaluación del programa.

- **Leer en el cuadro el tamaño de muestra necesario.** De acuerdo al cuadro, si  $P_1$  tiene un valor de 0,30 y  $P_2$  de 0,45, se requiere un tamaño de muestra de 353 jóvenes para cada ronda de recolección de datos.

*Nota:* Recordar que cada ronda de recolección de datos requiere este tamaño de muestra. En el ejemplo anterior, se necesitaría una muestra de 353 jóvenes en la encuesta base y otros 353 en la encuesta de seguimiento.

### **Muchos evaluadores optaron por agregar un margen para compensar por la deserción y las omisiones**

En todos los estudios, algunos sujetos no cooperan o abandonan el estudio. Este problema generalmente se aborda aumentando el tamaño de la muestra necesaria en una proporción fija (alrededor del 10%). Si el cuadro anterior recomienda una muestra de 843, es conveniente aumentarla un 10%, a 927, para compensar por los que no contestan.

#### **Cálculo del tamaño de la muestra**

Supóngase que el objetivo del programa es reducir la proporción de alumnos varones que consideran aceptable que alumnos varones presionen a las alumnas a tener relaciones sexuales. Se estima que el 50% de los alumnos varones sentían que este comportamiento era aceptable al comenzar el programa ( $P_1 = 0,50$ ). El programa aspira a reducir este porcentaje por lo menos 10 puntos porcentuales ( $P_2=0,40$ ). Si se usaran niveles de significancia y poder de 90%, el tamaño de la muestra requerida sería de 843 estudiantes varones para cada ronda de recolección de datos.

*Nota:* Este ejemplo ilustra el punto que el tamaño de la muestra requerida para medir un cambio de 0,50 a 0,40 es igual a la que se requiere para medir un cambio entre 0,40 y 0,50. Esto se aplica a cambios equivalentes en otros niveles.

Tasas de deserción de 25 a 30% o más son comunes en estudios de seguimiento de jóvenes. Para continuar con el ejemplo anterior, después de agregar un margen de 10% por los que no contestan, se debe considerar agregar un 25% adicional como precaución por los desertores, elevando la muestra base de 927 a 1.159. Esto ayudará a asegurar que se podrán detectar cambios en los indicadores, si estos ocurren, aun cuando se tiene desertores y personas que no contestan.

## **Para escenarios que no están cubiertos en el cuadro de tamaño de la muestra**

En el Anexo 2 hay un cuadro más completo que muestra los requisitos de tamaño de la muestra para diferentes combinaciones de niveles de significancia y poder. Ese Anexo también incluye una fórmula para calcular el tamaño de la muestra requerida y pautas para elegir los parámetros a usar en las fórmulas.

### **Calcular el tamaño de la muestra para medir cambios en los indicadores que involucran más de un comportamiento requiere un paso adicional**

Cabe notar que para algunos indicadores—especialmente los que involucran más de un comportamiento—se necesita un paso adicional para calcular el tamaño de la muestra. Ilustraremos este método usando el ejemplo del indicador: “Proporción de jóvenes que usaron un preservativo durante su última relación sexual”. Este indicador requiere medir dos comportamientos: (1) la proporción de jóvenes que tuvieron relaciones sexuales y (2) la proporción de jóvenes que usaron un preservativo durante la relación.

### **Calcular el tamaño de la muestra que se necesita para medir la magnitud del cambio en el uso de preservativos**

Usando el cuadro de tamaño de la muestra, determinar cuántos jóvenes se deberían encuestar para medir un cambio en la proporción que usó un preservativo durante su última relación sexual. Si se estimó que el uso de preservativos durante la última relación sexual fue de 30% ( $P_1$ ) y se espera aumentar esa proporción a 45% ( $P_2$ ), entonces se necesita tener una muestra de 353 jóvenes.

### **Calcular cuántos jóvenes se necesitará encuestar para obtener el tamaño de muestra necesario de jóvenes sexualmente activos**

Este indicador sólo puede medirse con jóvenes que son sexualmente activos, como por ejemplo aquellos que han tenido relaciones sexuales durante los últimos seis meses. Suponiendo que no todos los jóvenes son sexualmente activos, se necesita estimar cuántos jóvenes habría que entrevistar para encontrar 353 que sean sexualmente activos.

En primer lugar, hay que estimar la proporción esperada de jóvenes que han tenido relaciones sexuales en los últimos seis meses; por ejemplo, se puede pensar que 40% de los jóvenes en la población objetivo son sexualmente activos. Entonces se divide el tamaño de muestra requerido (353) por la proporción esperada de entrevistados elegibles (0,40). El número de jóvenes que se necesita entrevistar para encontrar 353 jóvenes sexualmente activos es  $(353 / 0,40)$ , así es que  $n=883$  para medir el indicador relacionado al uso de preservativos en la última relación sexual.

## Calcular el tamaño de la muestra para medir cambios en dos conductas

Supongamos que el objetivo del programa es aumentar la proporción de alumnos de la escuela secundaria que usan un preservativo durante su primera relación sexual.

### ***Calcular el tamaño de la muestra necesaria para medir la magnitud del cambio en el uso de preservativos.***

Se piensa que actualmente sólo alrededor del 10% de los estudiantes usan un preservativo en su primera relación sexual ( $P_1 = 0,10$ ). El programa quiere aumentar ese porcentaje a 20% por lo menos ( $P_2 = 0,20$ ). Con un 90% de niveles de significancia y poder, el tamaño de la muestra requerida de acuerdo al cuadro es  $n = 438$  estudiantes por cada ronda de recolección de datos.

### ***Calcular cuántos jóvenes se necesitará encuestar para obtener el tamaño de la muestra necesaria de jóvenes que han tenido una primera experiencia sexual.***

Las estimaciones actuales son que entre el 10 y 15% de los estudiantes tienen en una primera relación sexual cada año. Por lo tanto, en un periodo de tres años de un proyecto, se espera que entre 30 y 45% de los estudiantes tengan su primera relación sexual. Como un 30% de los estudiantes tendrá su primera relación sexual durante el periodo de tres años de un proyecto, el tamaño de la muestra necesaria por cada ronda de recolección de datos es  $n = 438 / 0,30 = 1.460$  estudiantes. Para detectar un cambio de 15 puntos porcentuales en el uso de preservativos, se requiere  $n = 720$  estudiantes por cada ronda de recolección de datos ( $n = 216$  del cuadro de tamaño de la muestra / 0,30).

La parte difícil de calcular el tamaño de la muestra para indicadores que involucran más de un comportamiento es estimar cuál es la proporción subyacente apropiada, por ejemplo, estimar la proporción de jóvenes que son sexualmente activos. Los estimados se pueden basar en otras encuestas o en informaciones anecdóticas del terreno. Sin embargo, si los recursos lo permiten, hacer una pequeña encuesta piloto puede ayudar a estimar esta proporción.

Si no se está seguro acerca de un estimado para un comportamiento específico, es mejor errar en el lado de la *subestimación* de la proporción involucrada en el comportamiento específico. Mientras más se subestime el comportamiento específico en la población objetivo, más alto será el tamaño resultante de la muestra. Por ejemplo, si no se está seguro si el 40 o el 50% de los estudiantes en la población objetivo tienen relaciones sexuales en un periodo de seis meses, se debería usar la cifra de 40% para determinar los requisitos de tamaño de la muestra.

## **Los tamaños de muestra necesarios para medir cambios en indicadores para unidades agregadas son menores**

Muchos proyectos pueden querer medir cambios en los indicadores a nivel de escuelas, servicios o comunidades. Los requisitos de tamaño de la muestra para medir estos indicadores son *la mitad del tamaño* indicado en la tabla. La razón para esto tiene que ver con los efectos del diseño en el muestreo por conglomerados. Los efectos del diseño se tratan en el Anexo 1.

Por ejemplo, supongamos que se quieren medir los cambios en la proporción de escuelas que han implementado programas educativos de habilidades para la vida que cumplen con ciertas pautas. Si pensamos que sólo el 10% de las escuelas tenían ese tipo de programa al comienzo del periodo de observación y queremos medir un cambio de 15 puntos porcentuales, el cuadro de tamaño de la muestra dice que se necesitan 165 escuelas en la muestra. Debido a que se está haciendo un muestreo de una unidad agregada, se necesita una muestra de sólo la mitad de ese número—es decir, sólo 83 escuelas ( $83 = 165/2$ ).

### Si el tamaño de la muestra excede el número de "localidades" disponibles

Algunos programas pueden calcular el tamaño de la muestra requerida para evaluar un programa basado en la escuela y encontrar que el tamaño de la muestra recomendada es mayor que el número total de escuelas cubiertas por el programa. En este caso, la evaluación debe incluir el máximo de localidades posibles.

### *...Preguntas frecuentes acerca del muestreo*

#### El programa tiene muchos indicadores que se medirán a través del mismo esfuerzo de recolección de datos. ¿Qué indicadores se deben utilizar para estimar el tamaño de la muestra?

Idealmente, se debe calcular el requisito de tamaño de la muestra para cada indicador y elegir el mayor de estos como el tamaño de muestra a ser usado. Esto asegura que se satisfagan los requisitos para todos los indicadores. Sin embargo, si estamos trabajando con un gran número de indicadores, esto puede ser tedioso. En este caso, podemos:

- *seleccionar algunos de los indicadores más importantes*
- *calcular los requisitos de tamaño de la muestra para esos indicadores*
- *usar el resultado más alto de estos cálculos para determinar el tamaño de la muestra*

Usar esta estrategia tiene dos desventajas: (1) los cambios en algunos indicadores pueden medirse más precisamente de lo necesario y (2) los cambios en otros indicadores pueden no medirse con la suficiente precisión.

#### ¿Qué pasa si se quiere comparar diferentes subgrupos de jóvenes, escuelas o servicios?

Comparar cambios en indicadores para diferentes subgrupos requiere una muestra de mayor tamaño. Por ejemplo, si queremos medir cambios en los indicadores para jóvenes varones y mujeres por separado, necesitamos calcular el tamaño de la muestra separadamente para cada género.

Por ejemplo, volvamos al objetivo del programa "aumentar la proporción de alumnos de la escuela secundaria que pretenden usar un preservativo durante su primera relación sexual". Si se estimó que el 20% de los jóvenes varones tiene la intención de usar un preservativo durante su primera experiencia sexual ( $P_1$ ), y se quiere elevar esta cifra en 15 puntos, a 35% ( $P_2$ ), se necesita una muestra de 300 hombres jóvenes. Si se estima que el 30% de las mujeres tenía la intención de usar un preservativo durante su primera relación sexual ( $P_1$ ), y se quiere aumentar esto a 45% ( $P_2$ ), se necesita una muestra de 353 mujeres. Por lo tanto, se necesita una muestra de 653 jóvenes en total, 300 muchachos y 353 mujeres jóvenes.

| Cálculos del tamaño de la muestra |       |       |     |
|-----------------------------------|-------|-------|-----|
|                                   | $P_1$ | $P_2$ | n   |
| <b>Hombres jóvenes</b>            | 0,20  | 0,35  | 300 |
| <b>Mujeres jóvenes</b>            | 0,30  | 0,45  | 353 |

### **¿Qué pasa si se planea efectuar más de dos rondas de recolección de datos?**

Los tamaños de la muestra incluidos en la sección anterior tienen que ver con la magnitud de cambios de los indicadores que sería detectada comparando dos rondas de recolección de datos. Si se planea efectuar más de dos rondas de recolección de datos, se pueden usar muestras más pequeñas por dos razones:

- *si el hecho de efectuar varias rondas de recolección de datos resulta en un período más largo de observación entre la ronda inicial y la final, habrá mayor probabilidad de que la magnitud de cambio en los indicadores sea mayor, requiriendo así muestras más pequeñas, y*
- *se pueden usar algunos procedimientos estadísticos que requieren muestras más pequeñas en el análisis (por ejemplo, métodos de medidas repetidas)*

Si se planean más de dos rondas de recolección de datos, se recomienda consultar con investigadores locales (por ejemplo, una universidad local) para decidir qué tamaño de muestra usar.

### **¿Se deben mantener las mismas localidades en cada ronda de recolección de datos?**

Desde el punto de vista estadístico, la ventaja de volver a la misma muestra de escuelas, servicios o comunidades en cada ronda de recolección de datos es que, generalmente, esto aumenta la precisión de la evaluación, sobre todo cuando se cubre un número de localidades relativamente pequeño. Sin embargo, existe una desventaja potencial: volver a la misma localidad puede causar que, gracias al previo aviso, los sitios elegidos implementen el programa aún más rigurosamente que los que no han sido elegidos para el estudio, pero que también participan en el programa. Esto puede resultar en resultados sesgados. Aunque se recomienda retener la misma muestra de localidades durante una evaluación, el administrador del programa debe asegurarse de que el programa no se ejecute con mayor rigor en las localidades escogidas para la evaluación que en las que no lo fueron.

### **¿Se debe mantener la misma muestra de jóvenes en cada ronda de recolección de datos?**

Las evaluaciones en las que se cubre a los mismos jóvenes en cada ronda de recolección de datos se llaman *estudios de panel*. Por la misma razón que mantener la misma muestra de escuelas, comunidades, etc. aumenta la precisión con que se miden los cambios, los estudios de panel son la mejor forma de medir cambios en comportamientos entre jóvenes. Sin embargo, los estudios de panel dependen de la retención y seguimiento de los mismos sujetos, lo que toma tiempo y es costoso. Por eso, los estudios de panel efectuados durante largos períodos (por ejemplo, dos o más años) se recomiendan sólo cuando el programa tiene una evaluación de largo plazo y un presupuesto substancial.

Sin embargo, cabe notar que a menudo es factible recolectar datos de los mismos individuos en cortos períodos de tiempo. Por ejemplo, muchas evaluaciones de programas basados en escuelas pueden seguir muestras de estudiantes durante períodos de uno o dos años.